Non-parametric Methods for Automatic Exposure Control, Radiometric Calibration and Dynamic Range Compression

Ahmed Sohaib

A thesis submitted for the degree of

Doctor of Philosophy at

The Australian National University

October 2018

© Ahmed Sohaib 2016

Except where otherwise indicated, this thesis is my own original work.

Ahmed Sohaib

24 October 2018

to my family.

Acknowledgments

I would like thank Almighty Allah for guiding me, and giving me strength to achieve this milestone. I am grateful to my primary supervisor, Dr. Antonio Robles-Kelly, for his energies, efforts, and guidance. I have always found him encouraging and motivating to work. His positive energy has driven my work to this end. I would also like to thank my co-supervisor, Dr. Nariman Habili, for his time to review my drafts and his concerned advices whenever needed. I would like to thank Dr. Hongdong Li for taking up the role of adviser and being part of my supervisory panel. His research contributions are inspiration to many.

I would like to thank The Australian National University and National ICT Australia for funding my PhD and providing a great learning and research environment. I am grateful to the staff and students of Canberra Research Lab at NICTA for their support and fellowship. I am also grateful to our Imaging Spectroscopy and Scene Analysis (ISSA) research team (Ran Wei, Cong Huynh, Lin Gu, Sejuti Rahman, Lawrence Mutimbu, Luis Romero-Ortega, and Arash Shahriari) for their wonderful support and discussions that paved my research path. I must thanks Khurrum Aftab, Adnan Shah, Saeed Anwar, Muhammad Hanif, Muhammad Usman, Zeeshan Hyder, Behrooz Nasihatkoon, and many others for their acquaintance and making my stay in Australia comfortable and enjoyable. I must thank the wonderful Canberra Community for making my stay in Canberra memorable and cherishing. It was a great learning experience living in a multi cultural place like Canberra.

In the end, I would like to thank my father, who passed away during my PhD, for his sacrifices and devotion to see my achievements. I would also like to thank my mother for her prayers and sacrifices for my success in every step of my life. I am also thankful to my sisters for their encouragement and prayers. Finally, I owe a lot to my wife, my son, and my daughter for bearing with me the difficulties of studentship abroad, and always there to comfort me with their joyful company.

Abstract

Imaging systems are essential to a wide range of modern day applications. With the continuous advancement in imaging systems, there is an on-going need to adapt and improve the imaging pipeline running inside the imaging systems.

In this thesis, methods are presented to improve the imaging pipeline of digital cameras. Here we present three methods to improve important phases of the imaging process, which are (i) "Automatic exposure adjustment" (ii) "Radiometric calibration" (iii) "High dynamic range compression". These contributions touch the initial, intermediate and final stages of imaging pipeline of digital cameras.

For exposure control, we propose two methods. The first makes use of CCD-based equations to formulate the exposure control problem. To estimate the exposure time, an initial image was acquired for each wavelength channel to which contrast adjustment techniques were applied. This helps to recover a reference cumulative distribution function of image brightness at each channel. The second method proposed for automatic exposure control is an iterative method applicable for a broad range of imaging systems. It uses spectral sensitivity functions such as the photopic response functions for the generation of a spectral power image of the captured scene. A target image is then generated using the spectral power image by applying histogram equalization. The exposure time is hence calculated iteratively by minimizing the squared difference between target and the current spectral power image. Here we further analyze the method by performing its stability and controllability analysis using a state space representation used in control theory. The applicability of the proposed method for exposure time calculation was shown on real world scenes using cameras with varying architectures. Radiometric calibration is the estimate of the non-linear mapping of the input radiance map to the output brightness values. The radiometric mapping is represented by the camera response function with which the radiance map of the scene is estimated. Our radiometric calibration method employs an L1 cost function by taking advantage of Weisfeld optimization scheme. The

proposed calibration works with multiple input images of the scene with varying exposure. It can also perform calibration using a single input with few constraints. The proposed method outperforms, quantitatively and qualitatively, various alternative methods found in the literature of radiometric calibration.

Finally, to realistically represent the estimated radiance maps on low dynamic range display (LDR) devices, we propose a method for dynamic range compression. Radiance maps generally have higher dynamic range (HDR) as compared to the widely used display devices. Thus, for display purposes, dynamic range compression is required on HDR images. Our proposed method generates few LDR images from the HDR radiance map by clipping its values at different exposures. Using contrast information of each LDR image generated, the method uses an energy minimization approach to estimate the probability map of each LDR image. These probability maps are then used as label set to form final compressed dynamic range image for the display device. The results of our method were compared qualitatively and quantitatively with those produced by widely cited and professionally used methods.

Contents

List of Figures

a

List of Tables

Chapter 1

Introduction

Imaging systems are central to modern day life. Ranging from space exploration (Martin et al. [2005]) to microscopic level studies (Lang et al. [2002]) imaging systems have a vital role to play. Imaging systems are also widely used today in hand held devices (Rohs and Gfeller [2004]) as well as in entertainment (Zhang [2012]) and mass media (Kubota et al. [1996]). Medical imaging has recently evolved as an important domain for non invasive diagnosis and life sciences studies (Groezinger [2000]). Remote sensing is a domain where spectral imagers are used for mining (Riley and Hecker [2013]), agriculture (Thenkabail et al. [2016]), surveillance (Pajares [2015]) and environmental (Nagendra et al. [2013]) applications. Spectral imagers are an important tool for material studies in material sciences (McCrindle et al. [2014]). Imaging systems are used in system automation (Yang et al. [2014]) and in the domain of robotics (Nagatani et al. [2013]) to perform extreme tasks. They also play an important role in quality assurance in food (Wu and Sun [2013]) and industry (Nevalainen and Pellinen [2016]). With the evolution of new technologies, there is a rise of a wide spectrum of applications which demand update and upgrade in imaging systems under use. Thus there is a continuous need to improve the imaging systems pipeline. In this thesis methods have been proposed to improve the imaging pipeline for a broad range of imaging systems including traditional RGB cameras as well as modern day multispectral and hyperspectral cameras. These include exposure control, radiometric calibration and dynamic range compression for low dynamic range display devices. To better understand the thesis contribution towards improving the imaging pipeline, we show, in Figure 1.1, an abstraction of the image acquisition pipeline presented by (Debevec and Malik [1997]).

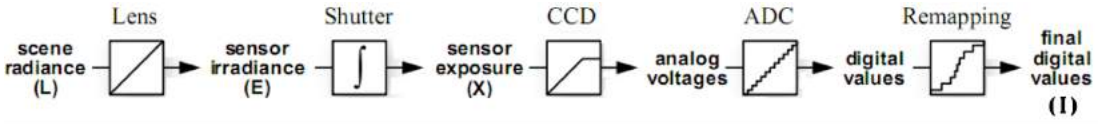

Figure 1.1: Abstract level Image acquisition pipeline of digital cameras proposed by (Debevec and Malik [1997])

At the abstract level, part of scene radiance passes through the lens, the aperture, and some desired filters to reach the sensor. At this stage, the imaging sensor (charged couple device CCD in this case) is exposed to the incoming light for a given exposure time (shutter speed) to generate photoelectrons. These photoelectrons are then converted to digital values using an analog to digital converter. After that, few on-board operations (e.g. noise reduction, sharpening, demosaicing, white balance, tone-mapping) are performed on the digital values. Finally, pixel values are obtained that are nonlinear with respect to the initial sensor irradiance.

We note that we can consider imaging devices as a system with input and output. We can control our input of light using the exposure adjustment method, and the resultant images are obtained as output. Our first contribution here is an automatic exposure control method for refining the input data to the imaging system. This is shown in Figure 1.2. Our method automatically adjust the exposure according to input scene such that maximum level of detail and contrast is obtained.

Exposure control plays an important role in determining the quality of the output of imaging systems. The importance of exposure control centers around the fact that

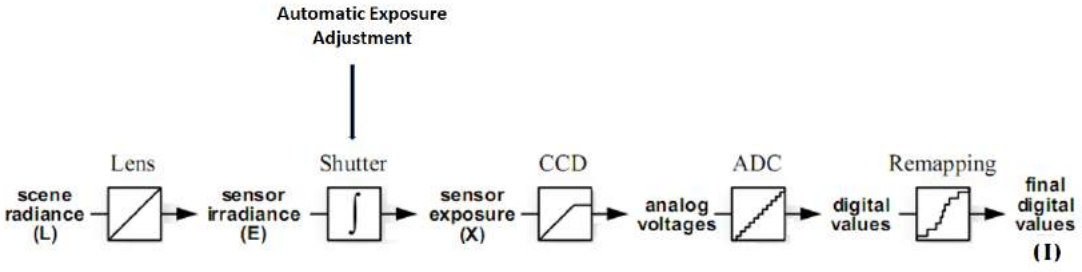

Figure 1.2: Our first contribution in proposing automatic exposure adjustment for cameras

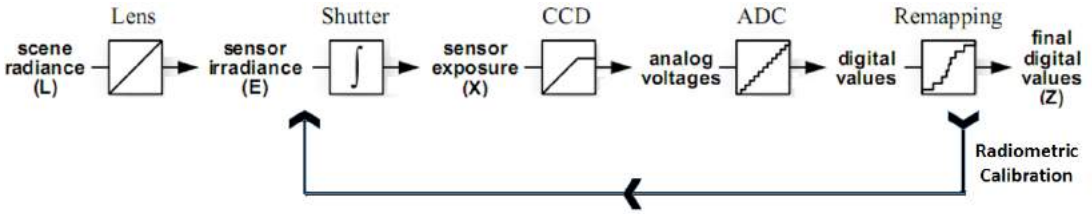

Figure 1.3: Our second contribution in proposing radiometric calibration of cameras

the dynamic range of cameras is extremely limited as compared to the dynamic range of natural scenes. Interestingly, an incorrect exposure setting can result in too dark or too bright output images, no matter how balanced the original scene was. Exposure control can be understood by assuming the dynamic range of the camera as a window, which can slide over the dynamic range of the natural scene. Thus exposure control methods slide this window to a position where the desired range of scene information can be maximized. It is quite natural to observe that as new imaging technologies and systems are evolving, so are the needs to explore methods to automate their exposure control.

Our next contribution is to propose a method for radiometric calibration to estimate the camera response function which maps the output image pixels to the input radiance map up to a scale. In Figure 1.3 we show the radiometric calibration applied to the output final image values to estimate the camera response function in order to estimate the sensor irradiance, or often referred as radiance map, up to a scale. Radiometric calibration is used in many computer vision applications especially problems related to scene analysis. It is also widely used in high dynamic range (HDR) imaging. An HDR image is obtained by estimating the radiance map of the scene, which is the resultant of radiometric calibration. These radiance maps are desirable in many applications spanning from gaming to digital media and panorama creation. In order to realistically represent the real world information, radiance map estimation is a desired step in imaging pipeline. Tone mapping or dynamic range compression is applied on high dynamic range radiance maps to generate images as camera output with finer level of details and contrast.

Digital cameras often acquire images with a reduced dynamic range, and HDR

imaging is usually achieved by combining differently exposed low dynamic range images of the same subject matter. In this manner, the problem of recovering an HDR image becomes that of radiometric calibration i.e. obtaining the radiance map of the scene from a series of bracketed exposure images by estimating the camera response curve (Banterle et al. [2011]; Mitsunaga and Nayar [1999]).

HDR images contain higher order of image details, which often is practically difficult to display on widely used low dynamic range (LDR) display devices. A dynamic range compression or tone mapping technique is required to map the high dynamic range values to the range of display devices. Our final contribution in this thesis is to propose a method for dynamic range compression for LDR displays such that the finer details remain unaffected to human vision. In Figure 1.4, we show the the high dynamic range radiance map obtained after radiometric calibration is then compressed for display purpose.

As discussed earlier, our previous approach delivers a radiance map image which is a linear representation of scene irradiance. This radiance map has a higher dynamic range and cannot be displayed directly to the display devices due to their lower dynamic range. Thus the radiance map images generated cannot be directly displayed to the display devices unless the dynamic range is compressed. This compression for mapping higher dynamic range to lower one is often referred as tone mapping. A good dynamic range compression algorithm is one that preserves the finer details of the image. Widely used dynamic range compression methods suffer from loss, amplification or reversal of contrast. Our proposed method for dynamic range compression for digital displays has an overall better performance than the widely cited and professionally used methods.

1.1 Overview

The remaining chapters of the dissertation are organized as follows.

In Chapter 2, we present the terminologies and basic understanding of processes taking place in an imaging pipeline. We also present a literature review of work undertaken that relates to our proposed methods, i.e. exposure control, radiometric

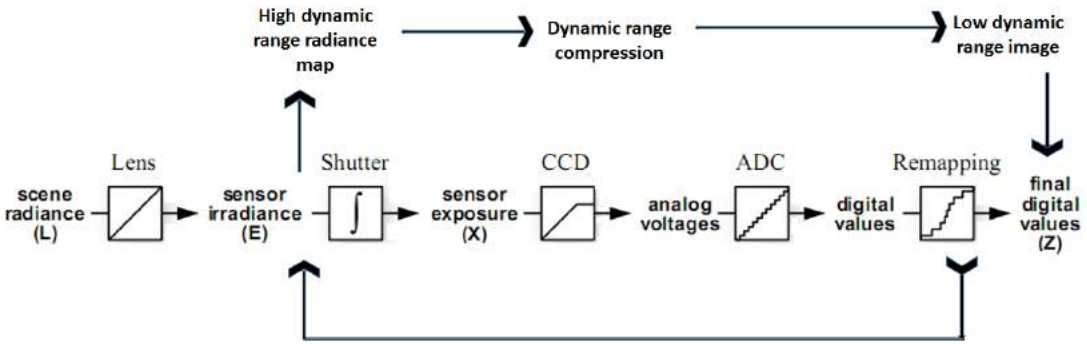

Figure 1.4: Our third contribution in proposing dynamic range compression of radiance map images for out images to digital displays

calibration, and dynamic compression.

In Chapter 3, we present our method for automatic exposure control of cameras, using histogram equalization technique.

In Chapter 4, another method of automatic exposure control is presented, which uses a spectral power image as a target image for estimating exposure time using optimization setting. In this chapter a detailed control and stability analysis is presented using state space representation.

In Chapter 5, our method to perform radiometric calibration technique is presented, which uses Weiszfeld L1 cost function for an accurate estimate of camera response function and radiance image.

Chater 6 presents our method for dynamic range compression. This chapter highlights our approach of exposure selection used for generating multiple LDR images, used in dynamic range compression.

Finally in chapter 7, the thesis is concluded with the avenues of future work.

1.2 Contributions

1.2.1 Automatic Exposure Control for Spectral Cameras Using Histogram Equalization

In chapter 3, we present a simple method to adjust the exposure of a hyperspectral camera. As hyperspectral cameras capture images over multiple wavelength-indexed

bands, exposure control techniques used for tri-chromatic sensors do not often extend to hyperspectral imagers. Here, we aim at estimating the exposure time for every wavelength channel by capturing an initial view of the scene. We then apply contrast adjustment techniques to this view of the scene so as to recover a reference cumulative distribution function (CDF) of the image brightness at each band. This CDF can then be used to calculate the new exposure time for every wavelength channel. We illustrate the utility of our method for exposure control and explore its stability with respect to both, a uniform and a Gaussian cumulative distribution functions.

1.2.2 Automatic Exposure Control For Cameras Using Optimization

Chapter 4 presents a method for automatic exposure time adjustment for cameras of different architectures such as staring array, multi-CCD, or traditional RGB cameras. The method presented here is based upon a spectral power image. Here, we use the photopic response function due to its widespread usage in photography and psychophysics. Note that, however, the method presented here is quite general in nature and can employ a number of spectral sensitivity functions for the computation of the spectral power image. Making use of this spectral power image, the exposure time is then computed via iterative updates so as to minimize the squared error between a target image and the current spectral power yielded by the imager. This target image is recovered, in a straightforward manner using histogram equalization and the CIE Photopic response function. This, in turn yields an automatic method devoid of calibration targets or additional inputs. We perform a stability and controllability analysis of our method using a state space representation. We also show the applicability of the method for exposure time calculation on staring array, multi-CCD, and single CCD architecture cameras on real-world scenes.

1.2.3 Radiometric Calibration of Cameras

In chapter 5, we present a method devoid of free parameters for radiometric calibration in which the radiance map and the camera response function is estimated for

high dynamic range (HDR) imaging. Our proposed method in its general form takes multiple images to reliably recover HDR images with as few as two low-dynamic range (LDR) images acquired at different exposures. As a special case our method can also be used to find camera response curve using a single well exposed image. To do this, we employ a L1 cost function which lends itself to the use of a Weiszfeld optimization scheme. We illustrate the utility of the method for computing high-dynamic range images from low-dynamic range imagery. In our experiments, we compare our results with those delivered by alternatives elsewhere in the literature and show that our method can outperform the alternatives with as few as two LDR images. We also apply our method to panorama generation.

1.2.4 Compressing Dynamic Range of Images for display

Chapter 6 presents a method for dynamic range compression of high dynamic range images. In the first step, our method generates a number of low dynamic range images at different exposures using the given high dynamic range image. It then uses the contrast taken from image luminance to generate probability maps. These probability map estimation is based upon an energy minimization approach that aims at recovering a set of labels for fusing the low dynamic range images. This treatment yields a convex minimization approach where the label-set can be computed using sparse Cholesky factorization techniques. We illustrate the utility of our method by generating the compressed dynamic range image and comparing qualitatively and quantitatively with the results of methods widely used in the literature and used professionally. Our quantitative comparison is based on three aspects of dynamic range compression, these are 1) loss of contrast 2) amplification of contrast, and 3) reversal of contrast.

Chapter 2

Background and Related Work

In this thesis we work on three parts of the imaging pipeline: (i) exposure adjustment at the time of image acquisition; (ii) radiometric calibration for accurate translation of scene radiance to image intensities; and (iii) dynamic range compression to display the translated radiance image to low dynamic range displays. Before explaning the proposed approaches, we present the basic working of CCD, based on which we shall build the proposed methods of imaging pipleline. In this chapter we also present a review of the related work done for our proposed methods.

2.1 Related terminologies

In this section, we briefly explain the terminologies related to radiometery. In radiometery we deal with methods and techniques to measure the electromagnetic radiation, which includes the visible part of electrmagnetic spectrum. On the other hand, in photometery we deal with the interaction of light with the human eye, thus, the visible part of the spectrum is the prime matter of interest in photometery. Here we briefly explain basic terminologies used in radiometery, which stands an important subject of the thesis.

In radiometery, we consider the light as a form of energy of the electromagnetic waves. It is referred as "radiant energy" and under International Standard of units (SI), its unit is joule (J). The radiant energy when related to consumed time becomes radiant power. The radiant power or radiant flux is the total power of the electromagnetic radiation. It can be emitted from a source or arriving at a surface. It's SI unit is watt (W) which is joule per second ($\frac{J}{s}$). The radiant power per wavelength of

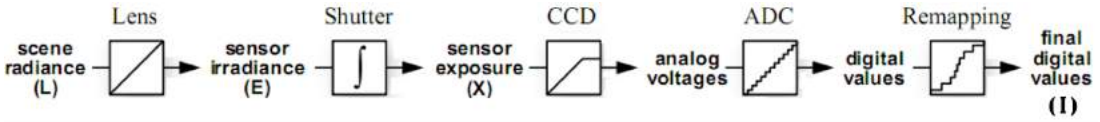

Figure 2.1: Abstract level Image acquisition pipeline of digital cameras proposed by Debevec and Malik [1997]

electromagnetic spectrum is referred as spectral power. Its SI unit is Watt per meter ($\frac{W}{m}$).

In radiometry, the often used term is radiance, which is power emitted or reflected by a surface and received by an optical system looking at the surface from the same angle of view. It defined as radiant flux in a given direction of a small element of surface area divided by the orthogonal projection of this area on to a plane at right angles to the direction of radiant flux. Its unit is $\frac{W}{m^{-2}sr^{-1}}$. Similarly, irradiance is the radiant power reaching the imaging sensor. Irradiance is the radiant power falling on a surface and is also referred as radiant flux density. Its SI unit is $\frac{W}{m^2}$. The sensor irradiance is comprised of radiant power of illuminant(s) reflecting from the object(s) in view with given reflectance(s). Reflectance of a surface is the ratio of the reflected energy of the incident light. Reflectance depends on the reflection geometry which includes the illuminant, viewing directions, surface normal, the power spectrum of the illuminant and the material properties such as roughness, albedo, shininess and subsurface structure (Robles-Kelly and Huynh [2012a]).

2.2 Background

In the previous chapter, an abstract level of image acquisition pipeline was shown. Here to further explain steps involved we reshow the image acquisition flow diagram in Figure 2.1. Each step in the flow diagram comprises many tasks. To begin, the light reaches the sensor after passing through the lens, aperture, shutter and any filter used. This can be graphically represented in Figure 2.2[1] where sensor irradiance contains radiant power passing through camera lens, aperture, and shutter. Assum-

[1]Image taken from 'Nikon D6000 for Dummies'

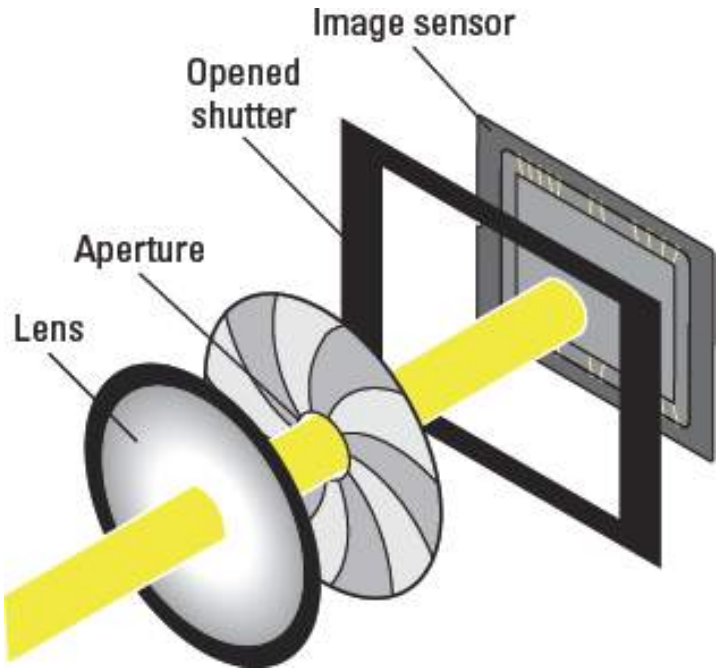

Figure 2.2: Graphical representation of light reaching imaging sensor through lens, aperture, with a desired shutter speed.

ing the lens effects the incoming light linearly, given as o, α as aperture size and κ as the effect of any filters used, the spectral irradiance reaching the sensor is given as

$$E = Lo\kappa\alpha \ . \tag{2.1}$$

Here L represents the scene irradiance. Note that these values are considered per wavelength channel and are integrated across the electromagnetic spectrum for a combined effect. For trichromatic RGB cameras, often color filters are placed over sensors to form a color filter array. The Bayer pattern is widely used for the color filter array. Assuming the sensor size as constant and T being shutter speed (the time the sensor is exposed) the sensor exposure for a given wavelength is given as

$$X = ET \ . \tag{2.2}$$

Charged couple device (CCD) sensors release photoelectrons when exposed to light. Each pixel location of the CCD contains a "*well*" where these photoelectrons are stored. The upper bound of photoelectrons well is called the full well capacity of

the pixel (Janesick [2001]). In case of higher exposure, pixel full well capacity often reaches its maximum, resulting in spill of photoelectrons to the neighboring pixels which causes image saturation and blooming effects (Klinger [2003]). After being stored in pixel wells, these photoelectrons are then counted and given a digital number called count for the given pixel location. The ratio between the incoming photons to the photoelectrons released by the CCD is known as quantum efficiency, whereas the conversion ratio of the photoelectrons number to a digital count is referred as gain of the CCD. Theoretically, if the shutter is closed and we capture an image with a short exposure time, there should not be any photons arriving at the CCD. However, in practice, the count is not null. This is due to the presence of dark current and CCD bias. Dark current is the small electric current flowing through the CCD when theoretically no photons are entering the camera. Dark current is dependent on the temperature of the CCD, where a high CCD temperature will result in higher dark current values. Bias often appears as a regular pattern on the image which arises from the pixel-to-pixel variations of the offset level on the CCD count. The dark current and bias should be subtracted from the image in order to properly estimate the photons and their conversion. Using the quantum efficiency, dark current and bias, we can express the count or data number as

$$O_{DN} = \frac{1}{\gamma}(EQT + \psi T + \beta) \ , \tag{2.3}$$

where O_{DN} is the observed data number or count, γ is the gain, E is the irradiance impinging on the CCD, Q is the quantum efficiency, T is the exposure time, ψ is the dark current and β is the bias.

To remove the effect of dark current and bias, the common practice is to take an image with a closed shutter, such that the photon flux becomes zero, *i.e.*, $EQT \to 0$. The new equation becomes

$$O' = \frac{1}{\gamma}(\psi T + \beta) \ . \tag{2.4}$$

If we subtract O' from O_{DN} our observed data number will be

$$O = \frac{1}{\gamma}(EQT) \ . \tag{2.5}$$

The expression above implies that the effects of dark current and bias can be safely ignored and removed automatically at start-up or with a periodic dark current acquisition routine. Note that this subtraction also assumes that that the effect of dark current and bias remains at the linear part of the camera response function.

The pixel values obtained are usually non-linear with respect to the incoming irradiance. The limited dynamic range of cameras pushes them to adjust the exposure according to the dynamic range of the scene. The pixel well capacity at saturation state adds to the nonlinearity of the camera response. Similarly, camera manufactures apply photo-finishing before saving the image in the final pixel values, (Nguyen et al. [2014]), which adds to the non-linearity of the final image values. This non-linear mapping is referred to as the camera response function. With this camera response function, the final pixel values is given as:

$$I = f\left(\frac{EQT}{\gamma}\right) . \tag{2.6}$$

These equations form the foundation of problem solving for exposure control, radiometric calibration and dynamic range compression. Depending on the nature of the problem, lens effect, aperture size, quantum efficiency, gain and filter effects are taken constant, and at times ignored, when the desired estimate is assumed up to a scale or normalized.

2.3 Related work

In this section, we present related work done for each of our contribution i.e. automatic exposure control, radiometric calibration and dynamic range compression.

2.3.1 Exposure Control for Cameras

Approaches for automatic exposure control in cameras can be classified into two classes, *Optics* and *Electronics* based. In optics-based approaches, light is measured by an independent sensor and the exposure is adjusted accordingly by changing the aperture or shutter speed of the camera. On the other hand, in electronics-based

approaches, the light is metered by the imaging sensor itself and the exposure is adjusted by changing the integration time and gain (Kremens et al. [1999]).

Early techniques often relied upon external hardware for light metering (Norwood [1940]). These metering devices often employed cadmium sulphide or silicon photo diode detectors (Hirokazu and Tadazumi [1975]). Later on, TTL (Through The Lens) mehtods brought the metering set-up inside the digital camera (Nelson [1966]). TTL metering techniques were widely employed in camcorders, their main drawback stemmed upon their computational complexity. Subsequently, photometric sensors have steadily replaced TTL systems for exposure control (Kuno et al. [1998]). These employ brightness information to estimate the exposure control parameters.

It is worth noting that, in the domain of auto-exposure control, most of the work is aimed at trichromatic cameras and disclosed in patents (US patent classification Class number 354 and SubClass 410-455). Patents like those in Bell et al. [2002]; Takagi [1997]; Johnson [1984] present a number of different methods to improve exposure control for trichromatic cameras. Bell et al. [2002], determines the initial exposure through a selection scheme applied to a set of predetermined exposure settings. This initial exposure time is employed to acquire an image, which is then used to find over or under exposed regions. These regions are then used to recover a new exposure setting via a look up table. These steps are performed iteratively until a final exposure setting is realised. In a similar development, Kuno et al. [1998] used the brightness values of the CCD and lookup tables in an iterative fashion in order to estimate the exposure time efficiently. Shimizu et al. [1992] has developed an algorithm for Sony video cameras which employs histograms and fuzzy logic techniques to determine the so-called front-lit and back-lit parameters, based on which the aperture size of the lens and the gain of the CCD can be determined.

Most modern automatic exposure algorithms use scene analysis techniques such as spot, matrix, evaluative and center-weighted metering to analyze the brightness of a scene or an object of interest and adjust the exposure setting accordingly. To compensate for back lit scenes, Lee et al. [2001] proposes a division of the image into several areas, assuming that the main object of interest tends to be in the center of the scene. Nourani-Vatani and Roberts [2007] presents a method where various

masks are applied so as to obtain the exposure adjustment using the settings for the lens, environment and the area of interest. Similarly Rogers and Cope [2010] uses human flesh regions in the scene to control the exposure of the camera. His work is particularly well suited to improve the appearance of the human subjects in portraits and snapshots acquired by hand-held devices.

Spectral cameras, unlike trichromatic cameras, provide information over a large number of wavelength channels across the electromagnetic spectrum. This, in effect, delivers an information-rich representation of the scene which can be used in areas such as detection (Dale et al. [2013]; Akbari et al. [2011]; Kulasekara et al. [2014]), classification (Caras and Karnieli [2015]; Garrido-Novell et al. [2012]) and recognition (Khan et al. [2011]; Long and Li [2011]; Zhao and Liu [2010]). Spectral imaging has also found application in areas such as colour constancy (Wandell [1987]) and the optimal multiplexing of bandpass filtered illumination (Schechner et al. [2003]; Chi et al. [2010]).

For hyperspectral cameras, calibration targets are often used to adjust the exposure time. An example for such calibration process is presented in the work of Brelstaff et al. [1995] and Pichette et al. [2017]. Unfortunately, photometric exposure calibration with a white reference target is infeasible or impractical in many real-world settings. It is worth noting that the image acquisition by hyperspectral cameras can be a complex task as each channel often has to be adjusted accordingly (Dwight et al. [2018]). Even the modern hyperspectral systems rely on look tables with predefined exposure values (Brucalassi et al. [2018],Dogan et al. [2017]) than using automatic exposure calculation.This particularly applies to staring arrays and multi-CCD cameras but also extends to systems that acquire multiple hyperspectral images on airborne or mobile platforms.

2.3.2 Radiometric Calibration

Due to the complexity of the imaging pipeline with various electronic effects there exist a non linear relation between the image brightness and sensor irradiance. This non linear mapping is often referred as radiometric response or camera response function (CRF). Estimating this camera response function and the radiance map from

image intensities is usually referred as radiometric calibration. Radiometric calibration is helpful in linearizing the image data which solves as well as improves many computer vision related application.

One major area where radiometric calibration is widely used is high dynamic range (HDR) imaging. HDR imaging is obtained by estimating the radiance map of the scene, which is the resultant of radiometric calibration. These radiance maps are desirable in many applications spanning from gaming to digital media and panorama creation. HDR images or radiance maps are also closely related to tone mapping (Eilertsen et al. [2017]). For instance, Banterle et al. [2007] have used a low-dynamic range (LDR) image to compute the inverse tone mapping curve so as to increase the dynamic range of the LDR image under study.

Since digital cameras often acquire images with a reduced dynamic range, HDR imaging is usually achieved by combining differently exposed images of the same subject matter. In this manner, the problem of recovering an HDR image becomes that of radiometric calibration i.e. obtaining the radiance map of the scene from a series of bracketed exposure images by estimating the camera response curve (Banterle et al. [2011]; Mitsunaga and Nayar [1999]). While many methods work on the assumption of aligned images, methods like (Kalantari and Ramamoorthi [2017]) includes images alignment as a preprocessing step for recovering HDR image.

An alternative set of methods for HDR imaging aim at using specially designed sensors, constrained image acquisition settings (Ikeda [1998]) or specialised optics (Nayar and Mitsunaga [2000]). Similarly Machine learning techniques have also been proposed using Neural Networks () but they involve large datasets for training purposes. Other techniques avoid the computation of a radiance map by fusing the imagery based upon chroma or luminance information. Sung et al. [2007] proposed a tone reproduction system which takes at input two images acquired at different shutter speeds. This method employs a local adaptation mechanism which considers the colour statistics and imposes spatial constraints on the image lattice. Similarly Jung and Ho [2013] proposed a method for texture detail enhancement which employs two images which are assumed to be taken at different exposures, *i.e.* f-stops. In Jung and Ho [2013], Markov random fields (MRF) over the luminance values

are used to enhance an under exposed image so as to fuse two input images. This technique is somewhat related to other methods such as those in Sun et al. [2010] and Jinno and Okuda [2012], which use MRFs to estimate HDR images using occlusions, saturated regions and local displacement compensation. These approaches often build upon the methods developed by Debevec and Malik [1997] or Mitsunaga and Nayar [1999] to estimate the camera response function so as to generate radiance maps for the HDR estimation. Robertson et al. [2003], have used the camera response function for the computation of and HDR image.

Lin et al. [2004] proposed a different model for radiometric calibration. Their method uses a single image for radiometric calibration. It focuses on the edge information of the image where assumption is that the radiance values remains unchanged whereas with the change of color intensities that curve is approximated. Lin and Zhang [2005] advance the single image method and use grayscale information by employing a histogram of edges. These methods working on edge information are often found prone to high level of image noise. Matsushita and Lin [2007] use noise distribution for radiometric calibration. The assumption is that for radiance maps obtained after radiometric calibration are symmetric. Whereas the noise distribution of input LDR image is asymmetric. Thus, the asymmetric profiles of the data are used for radiometric calibration. Similar assumption is used in the method of (Aguerrebere et al. [2014]). These approaches of working with noise distributions are often complex and suffers largely in low lighting conditions.

Here we note that, despite different approaches for radiometric calibration and HDR imaging, those techniques that require multiple LDR images often are simpler and more reliable to estimate the camera response function CRF. Methods with multiple images have increased range of data points which results in better estimation of camera response function. The main drawback of these methods stem from the fact that acquiring large number of images increases time and computational order of these methods.

2.3.3 High Dynamic Range Compression

High Dynamic range (HDR) imaging (Reinhard et al. [2010]) is widely used in the domain of photography (Hasinoff et al. [2010]) , computer graphics (Johnson and Fairchild [2003]) and scientific imaging (Malbet et al. [1995]; DeBoer et al. [2009]). HDR images contains fine details of the scene which can be displayed on range of display devices with proper mapping and compressing. Sensor manufacturers have attempted to increase the dynamic range by either decreasing the noise floor or increasing the full well capacity of the sensors (Yadid-Pecht [1999]; Spivak et al. [2009]). Widely used approach for dynamic range increase for acquisition devices is to generate radiance maps by capturing images at several exposures (Debevec and Malik [2008]; Sohaib and Robles-Kelly [2015]; Mitsunaga and Nayar [1999]; Jinno and Okuda [2012]; Robertson et al. [2003]). These radiance maps are used for HDR images generation. For displaying digitally the HDR images, the digital displays available have limited dynamic range and often requires expensive solutions for increasing their dynamic range (Seetzen et al. [2004]; Lieb and Russell [2014]). Thus the HDR images generated cannot be directly displayed on LDR displays unless the dynamic range is compressed. This compression for mapping higher dynamic range to a lower one is often referred as tonemapping. A good dynamic range compression algorithm is the one that preserves the finer details of the image.

Various methods and techniques have been proposed for dynamic range compression. Initial work to display HDR images on the digital display can be referred back to the work of Tumblin and Rushmeier [1991, 1993]. They are generally categorized in two forms, global techniques and local techniques. Global techniques are based on the global variables of the image like luminance. They estimate a function $f(.)$ which when applied to image $I(x, y)$ gives transformed image as $I'(x, y) = f(I(x, y))$. Once the function is estimated for an image it is applied to all pixels across board without considering the pixels neighborhood. The advantage of these techniques is run time efficiency and method simplicity. Common issues with this approach is loss of contrast and color inversion. In local techniques the parameters mapping or function changes spatially for each pixel based on its neighborhood. Thus local techniques are

more complex than global ones. These techniques might introduce unrealistic effects like halo or ringing. Since these techniques of preserving local contrast finds its basis in human visual system where local contrast sensitivity is higher, it provides better performance.

In global techniques a range of methods developed initially were based on image histogram (Tumblin and Rushmeier [1993]; Ward [1994]; Ferwerda et al. [1996]; Larson et al. [1997]). Drago et al. [2003] used logarithmic compression based on psychophysics approaches in which logarithm bases were adjusted depending on image contrast. Similarly Reinhard and Devlin [2005] used photo-receptor physiological response to globally compress the dynamic range of the images. Kautz [2008] proposed a global and consistent tone reproduction method which applies tonemapping on a range of HDR images.

Chiu et al. [1993] proposed one of the initial works on locally adaptation techniques. Tumblin and Turk [1999] used image diffusion for gradient mapping using partial differential equation solver. Fattal et al. [2002] uses luminance gradients at different scales to identify large gradients to attenuate their magnitudes. It penalizes the larger gradients more heavily than the smaller ones, to compress drastic luminance changes while preserving finer details. Durand and Dorsey [2002] used gain maps to prevent halos with bilateral filtered described by Tomasi and Manduchi [1998].

Perceptual models based on human vision system (HVS) have also been proposed for dynamic range compression. Pattanaik et al. [1998, 2000] proposed a model using quantitative measurements from physiology, psychophysics and color science for tone reproduction. Similar comparisons are performed by Liu et al. [2011] on six different algorithms. Similarly, Ashikhmin [2002] uses HVS functionality to estimate local adaption luminance at each pixel and apply compression function to shift values in display range. Image details are added in final pass. In the approach of Sudhama et al. [2018] used large number of human human observers to judge the loss in compression of HDR images.

Chapter 3

Automatic Exposure Control for Spectral Cameras Using Histogram Equalization

Exposure control for cameras is important as it defines the quality of information obtained from the images. The real world perceived by the human visual system has a broad dynamic range. Dynamic range is often referred as the ratio between the intensity of the bright and the dark point. The broader the range between the brightest and the darkest intensity, the higher the dynamic range is. Digital cameras posses limited dynamic range. Exposure control systems are designed in the cameras to adjust the dynamic range of the camera according to the varying lighting situation. Automatic exposure control is a desirable feature in environment where the illumination is prone to change, delivering images with a consistent, reproducible distribution of brightness values. It is also desired since the dynamic range (the ratio between the maximum and minimum intensity values) of many real-world scenes far exceeds that of many cameras. As a result, an incorrect exposure setting yields dark or overexposed imagery with a reduced contrast. This can be observed in figure 3.1 where a single scene is captured with varying exposure times. The image gets saturated or darkened due to extreme exposures.

In this chapter, we present a method for automatic exposure control of spectral cameras. We observe that the spectral cameras, unlike trichromatic cameras, provide information over a large number of wavelength channels across the electromagnetic

Figure 3.1: Effect of exposure time changes on the image brightness. From left-to-right: Pseudo-colour images (hyperspectral images from Foster dataset (Foster et al. [2006]), rendered into colour using the CIE colour matching function in CIE [1932]) acquired with increasing exposure times. Note the image detail is lost at either extreme of the timings, whereby the imagery is underexposed or saturated.

spectrum. Over the last decade, the production cost of spectral cameras have declined (Robles-Kelly and Huynh [2012b], page number xiii). Initially used in satellites and airborne, now their use in ground based applications for detection (Dale et al. [2013]; Akbari et al. [2011]; Kulasekara et al. [2014]), classification (Caras and Karnieli [2015]; Garrido-Novell et al. [2012]) and recognition (Khan et al. [2011]; Long and Li [2011]; Zhao and Liu [2010]) is expanding. Spectral imaging has also found application in areas such as colour constancy (Wandell [1987]) and the optimal multiplexing of bandpass filtered illumination (Schechner et al. [2003]; Chi et al. [2010]).

Spectral cameras often have a complex architecture, such as staring array [1], line-scan [2], multi-ccd [3], snap shot[4]. For this reason, automatic exposure control methods pertaining to the widely used trichromatic cameras (i.e. RGB cameras) cannot be directly used for spectral cameras, whereas, the converse may be feasible. This converse is generally applicable for for the automatic exposure control methods of starring array, multi-CCD, and snapshot spectral cameras as their architecture design makes RGB cameras a special case. This relation is represented in Equation 4.4 in Chapter 4 where c represents the channels and for RGB cameras, c shall be three. Thus our proposed method for exposure control of spectral cameras, has natural extension to the traditional RGB cameras. In spectral cameras, each wavelength indexed band can be viewed as an intensity image. It is observed that the illuminant, reflectance and the

[1] http://www.optoknowledge.com/hyper-lctf-features.html

[2] https://www.ximea.com/en/products/hyperspectral-cameras-based-on-usb3-xispec/mq022hg-im-ls150-visnir

[3] http://www.fluxdata.com/products/fd-1665-ms7

[4] https://www.ximea.com/en/products/hyperspectral-cameras-based-on-usb3-xispec/mq022hg-im-sm4x4-vis

sensor sensitivity vary across the electromagnetic spectrum, thus each band in the image exhibits different photometric behavior. This can be observed in Figure 3.2, where we show three sample bands for a landscape image. Note that the combined effect of the illuminant and scene reflectance results in significant difference in the brightness across the three wavelength-indexed bands in the figure.

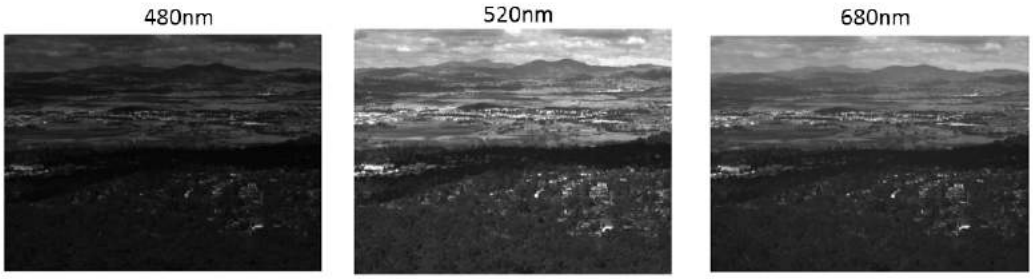

Figure 3.2: Sample wavelength indexed-bands for a hyperspectral image depicting a landscape. From left-to-right: Bands corresponding to 480nm, 520nm and 680nm.

In this chapter, we begin with basic equations of CCDs for spectral cameras (3.1). In Section 3.2, we explain the method of using histogram equalization for exposure time recovery in spectral cameras. In Section 3.3, we show the results of our method when staring array camera system based upon liquid crystal tunable filters. We show the results of our method test in varying lighting condition, with different histogram distributions, and varying initial exposure times. Finally in Section 3.4, we present the summary of the chapter.

3.1 CCD representation for spectral camera

Consider the spectral image I whose pixels are indexed to the row and column co-ordinates, and the wavelength index λ. In some cameras, every wavelength channel has an associated exposure time. This is often the case with staring array systems such as the OKSI hyperspectral [5] cameras. Other systems use a single exposure time for multiple wavelength indexed bands. This is the case for multiple CCD systems

[5] http://www.optoknowledge.com/hyper-lctf-features.html

such as the Fluxdata multispectral cameras [6]. In the case of the Fluxdata camera, seven wavelength channels are acquired from three CCDs. Each CCD has its own exposure setting. For the sake of generality, we assume that m wavelength channels are divided into c sets, where each wavelength set Λ_j has its own exposure time T_j for $j = 1, 2, ..., c$. For cameras where every wavelength has its own exposure time (e.g., the OKSI systems), c is equal to m. Let us assume that the pixel values we get from a camera are actual *data numbers*. We can then rewrite Eq. (3.1)

$$I = f\left(\frac{EQT}{\gamma}\right) , \tag{3.1}$$

as

$$I(v, \lambda) = E(v, \lambda)Q(\lambda)T_j , \tag{3.2}$$

where $j = 1, ..., c$, which relates to the wavelength sets. $I(v, \lambda)$ is the pixel value at location v and channel λ. $E(v, \lambda)$ is the Irradiance impinging on pixel v of channel λ, $Q(\lambda)$ is the quantum efficiency for the channel corresponding to the wavelength λ and T_j is the exposure time for the wavelength set Λ_j. Please note that the gain γ is a constant that has been absorbed into the equation above.

Further, we can write

$$E(v, \lambda)Q(\lambda) = \mathcal{V}(v, \lambda) , \tag{3.3}$$

where $\lambda \in \Lambda_j$. $\mathcal{V}(v, \lambda)$ is the combined effect of irradiance and quantum efficiency of the sensor at pixel location v of channel λ, which can be calculated using Equation 3.2. Note that due to higher quantization in spectral images, it is safe to assume a linear camera response function.

[6] http://www.fluxdata.com/products/fd-1665-ms7

3.2 Exposure Time Recovery Using Histogram Equalization Technique

Here we explain our method of exposure recovery using histogram equalization. For this we convert our equations in matrix representation. We write equation 3.2 as

$$\mathbf{O} = \mathbf{E}\mathbf{Q}T \ , \tag{3.4}$$

where $\mathbf{O}$ is a matrix of size $N \times c$ representing image intensity I arranged accordingly (in a row-column form), $\mathbf{E}$ is a $N \times c$ matrix containing the Irradiance values, $\mathbf{Q}$ is a $c \times c$ diagonal matrix whose diagonal entries are given by the quantum efficiencies for the CCD or CCDs in the camera and T is a $c \times 1$ vector representing the exposure times.

Since $\mathbf{E}$ and $\mathbf{Q}$ remain unchanged for a static scene, we can write

$$\mathbf{V} = \mathbf{E}\mathbf{Q} \ , \tag{3.5}$$

such that

$$\mathbf{O} = \mathbf{V}\mathbf{T} \ . \tag{3.6}$$

Note that we can calculate the matrix $\mathbf{V}$ directly using the Penrose pseudo inverse of T given by

$$\mathbf{V} = \mathbf{O}T^t(TT^t)^{-1} \ . \tag{3.7}$$

By acquiring an image of a scene with an initial exposure time, it is possible to enforce a histogram distribution upon the pixel values for each channel so as to recover the vector T of exposure times. The initial exposure time here shall be a vector where each member of the vector represents the exposure time of individual channel. The initial exposure time is preferred to have smaller values so that exposure is adjusted in minimum possible time. It is important that the initial exposure must not be set on extremes so as to avoid dark or washed initial image, where the no distinct information is available. With a reasonable initial exposure the initial image can be

used to compute a contrast-adjusted image by applying histogram equalization so as to recover an enhanced $\mathbf{O}_{new}$. Since $\mathbf{V}$ remains unchanged, the new exposure times T_{new} can be easily computed.

In literature (Kaur et al. [2011]) there are a number of alternatives that can be used for purpose of contrast equalization and enhancement. These "stretch" and transform the pixel values in the image so as to fit a target distribution. These target distributions can be a uniform one, stretched Gaussian, etc. Moreover, these methods are usually performed by obtaining the histogram of the image intensities so as to compute a Cumulative Distribution Function (CDF). Then, an inverse CDF of the desired target is applied to the original cumulative distribution as computed from the image histograms. For histogram equalization, the target distribution is set to be Uniform, which is the preferred target distribution for most of the cases, whereas other distributions may be required based on the scene information and user preference of information display.

Let f be the CDF of the initial image and g be that of the target. Similarly, let g^{-1} be the inverse of desired target CDF. The contrast enhancement step in our method can then be written as

$$\mathbf{O}_{new} = g^{-1} f(\mathbf{O}_{old}) \ . \tag{3.8}$$

Here $\mathbf{O}_{old}$ denotes the first image. The function f transforms this image into its cumulative distribution whereas g^{-1} is the inverse CDF for our desired target. Making use of Equation 3.6, we can rewrite Equation 3.8 as follows

$$T_{new}\mathbf{V} = g^{-1} f(T_{old}\mathbf{V}) \ . \tag{3.9}$$

where T_{old} is the exposure time set for the first image, whereas T_{new} is the exposure time set yielded by our method.

For a uniform distribution target and a probability p such that $0 \leq p \leq 1$ the inverse of CDF is given is as

$$g^{-1}(p) = \sigma\sqrt{3}(2p-1) + \mu \ . \tag{3.10}$$

Thus Equation 3.9 becomes

$$T_{new}\mathbf{V} = \sigma\sqrt{3}(2f(T_{old}\mathbf{V}) - 1) + \mu \tag{3.11}$$

where σ is the standard deviation and μ is the mean of the cumulative distribution for the first image.

Similarly, for a Gaussian distribution, Equation 3.9 is given as follows

$$T_{new}\mathbf{V} = \sigma\sqrt{2}\epsilon(2f(T_{old}\mathbf{V}) - 1) + \mu \tag{3.12}$$

where $\epsilon(\cdot)$ is the inverse of Gauss error function which can be defined in terms of a Maclaurin series as

$$\epsilon(z) = \sum_{k=0}^{\infty} \frac{c_k}{2k+1}(\frac{\sqrt{\pi}}{2}z)^{2k+1} \ . \tag{3.13}$$

In the equation above, we have set $c_0 = 1$ and

$$c_k = \sum_{m=0}^{k-1} \frac{c_m c_{k-1-m}}{(m+1)(2m+1)} \tag{3.14}$$

With the term $T_{new}\mathbf{V}$ available, T_{new} can be computed using the Penrose pseudoinverse of $\mathbf{V}$ in equation 3.9. This yields

$$T_{new} = (\mathbf{V}^t\mathbf{V})^{-1}\mathbf{V}^t g^{-1} f(T_{old}\mathbf{V}) \ . \tag{3.15}$$

3.3 Experiments

In this section, we illustrate the utility of our histogram equalization approach for the purposes of exposure time adjustment. For our experiments, we have used imagery acquired by OKSI staring array hyperspectral camera system based upon Liquid Crystal Tunable Filters (LCTFs). Our imagery was acquired across 31 wavelength-indexed bands in the visible spectrum at $10nm$ intervals in the range $[410nm, 710nm]$. The images were acquired using random initial exposure times between $50ms$ and $100ms$ and depict both, indoor and outdoor scenes. In all our figures, we show colour images, *i.e.* pseudocolour, computed from hyperspectral images using the CIE colour

Figure 3.3: From left-to-right: Pseudo color image acquired using random exposure times set, equalisation results for both, the uniform and Gaussian distributions and imagery acquired using the exposure times yielded by our method for the two equalisation targets.

matching functions in CIE [1932].

In figure 3.3, the left-hand panel shows the pseudo color image for the hyperspectral image acquired using randomly selected exposure times. In the second and third columns, we show the images obtained after the histogram equalisation with the uniform and Gaussian targets has been applied to the imagery on the left-hand column. In the fourth and fifth columns, we show the pseudocolour images for the imagery acquired using the exposure times yielded by our method. Figure 3.4 shows the initial and final exposure times as a function of wavelength index. In the plot, the initial exposure time is shown in red, whereas those corresponding to the Gaussian and uniform distribution targets are plotted in blue and green traces. From the plots, note that changing the target distribution does effect the recovered exposure times. This is a consequence of the difference in the contrast-adjusted images delivered by the two targets.

In figure 3.5, we show per-band results for an outdoor scene. In the figure, in each of the two panels (top and bottom), the columns correspond to a wavelength indexed band in the image. From top-to-bottom, these columns show the band acquired using a randomly selected exposure time, the equalisation results, the band captured using the exposure time yielded by our method, and the per pixel error map between equalisation and our method results. In the figure, we have used a uniform CDF target. Note that the pixel-wise errors are correlated with the band brightness. This is somewhat expected, since the larger the brightness for the equalised imagery, the more noticeable the discrepancies between the two images will be.

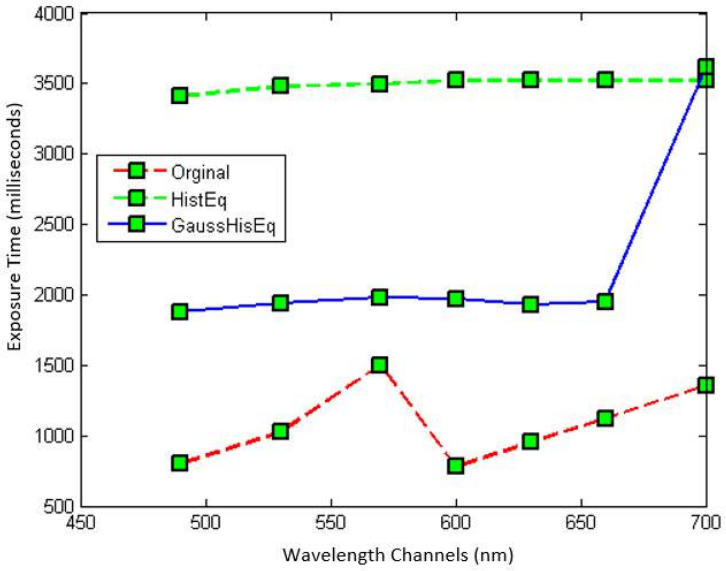

Figure 3.4: Exposure times as a function of wavelength index for the imagery in Figure 3.3. X-axis represents the wavelength channels in nm. Y axis represents the exposure time in milliseconds.

In Figure 3.6, we explored the effect of varying the initial exposure times used to acquire our images. In the figure, we show the results for 10 trials on the same scene. From left-to-right, each row in both panels shows the pseudo colour image acquired using a randomly selected exposure time followed by the image captured using the times yielded by our method and the time plots as a function of wavelength index. In the figure, we have used a uniform distribution target and, similarly to the plot in Figure 3.4 , the random exposure time trace is shown in red whereas the time yielded by our method is shown in green. From the plots, note that our method generates quite stable exposure times across different trials. The figure also hints to the fact that, for our method to yield sensible results, the initial images should not be overly under or over exposed. For these cases the computed exposure times are far removed from the expected ones. This can be seen in the 9^{th} image, where the exposure time of fourth channel is significantly different from those in other images. This is due to its randomly selected exposure time, which is quite low as compared to the others.

In Figure 3.7, we show the mean exposure times for the 10 trials in Figure 3.6. In the figure, the x-axis represents the wavelength-indexed bands and the y-axis ac-

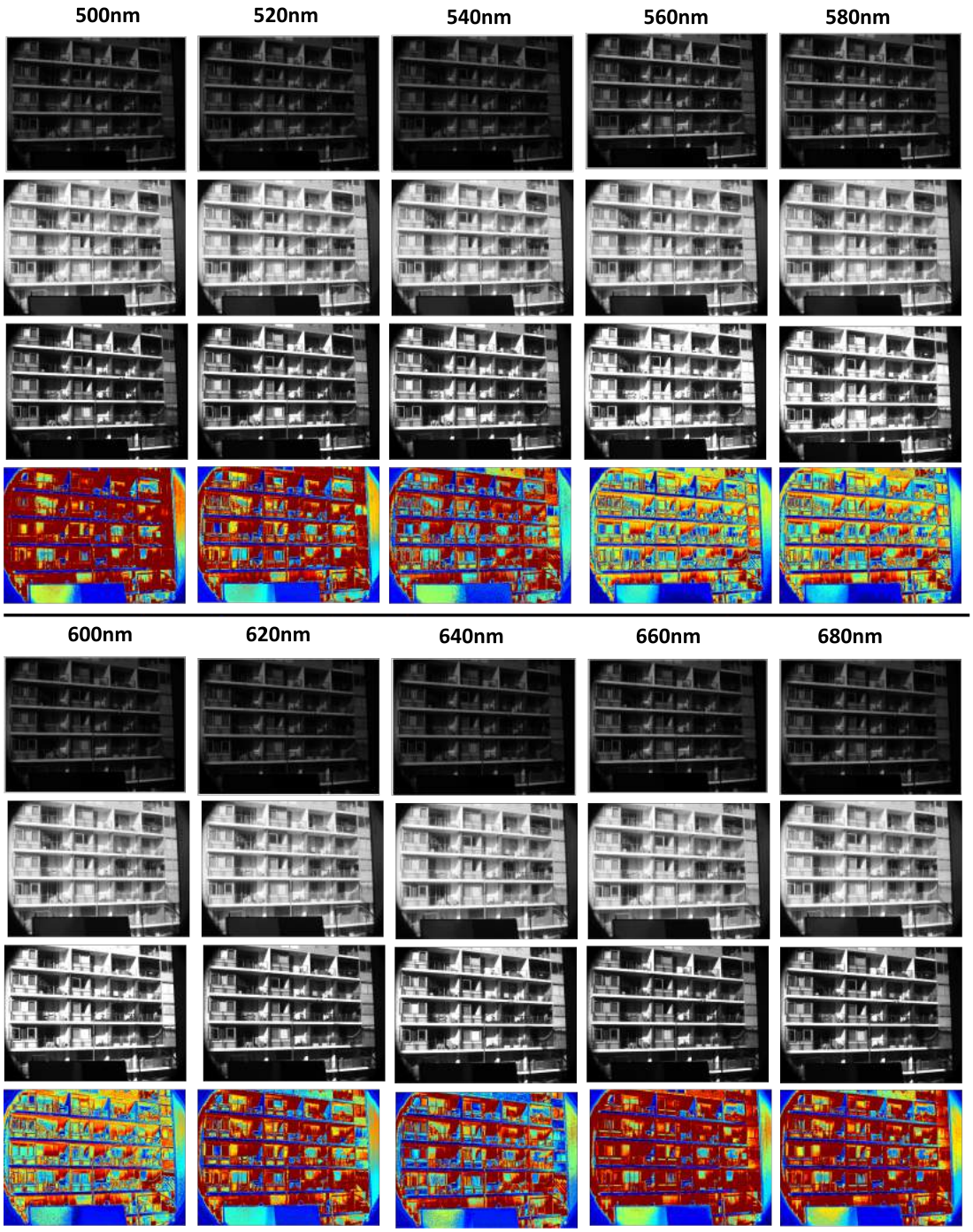

Figure 3.5: Results yielded by our method per wavelength indexed band. For both top and bottom panels, each column, from top-to-bottom shows: Band acquired using a randomly selected exposure time, histogram equalised image, wavelength-indexed band acquired using the exposure time yielded by our method and per-pixel normalised error map. For the error maps, red represents a value of 1 and blue amounts to zero.

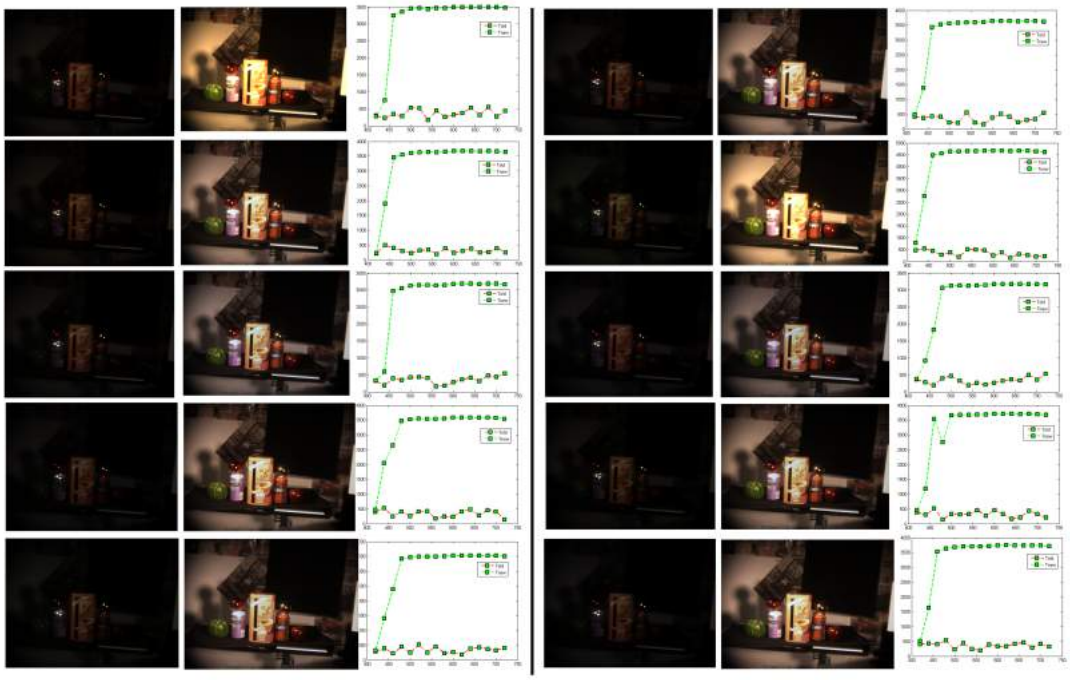

Figure 3.6: Results of our method for 10 trials corresponding to randomly selected exposure times as input to our method. For each of the two panels in the figure, each row corresponds to, from left-to-right, pseudo colour image for the images acquired using randomly selected exposure times, imagery captured using the times delivered by our method, time plots showing the initial and recovered exposure times as a function of wavelength index.

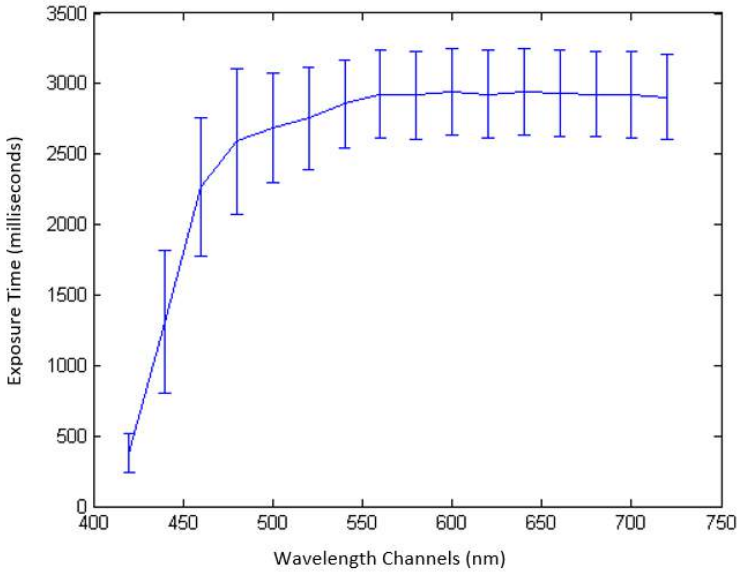

Figure 3.7: Mean exposure times for the 10 trials in Figure 3.6. X axis represents the wavelengths and Y axis represents the exposure time.

counts for the exposure time in miliseconds. In the plot, the error bars correspond to the standard deviation across trials. Note that the variance remains almost constant across the wavelength channels.

Finally, in Figure 3.8, we illustrate how our method adjusts itself with respect to varying light conditions and scenes. In the figure, in each row we show the pseudo colour images for the initial spectral image, followed by the one acquired using the exposure times yielded by our method and the corresponding time plots. Each row presents different lighting conditions. Note that the imagery acquired using the exposure times computed with our approach exhibit good contrast throughout. The plots also hint to the ability of the method to adjust itself with respect to changing lighting conditions. As before, the red plot corresponds to the random exposure times whereas the blue trace represents the exposure times yielded by our method.

3.4 Summary

In this chapter, we have presented a method for automatically recovering the exposure times for a staring array or multi-CCD camera system. Departing from the

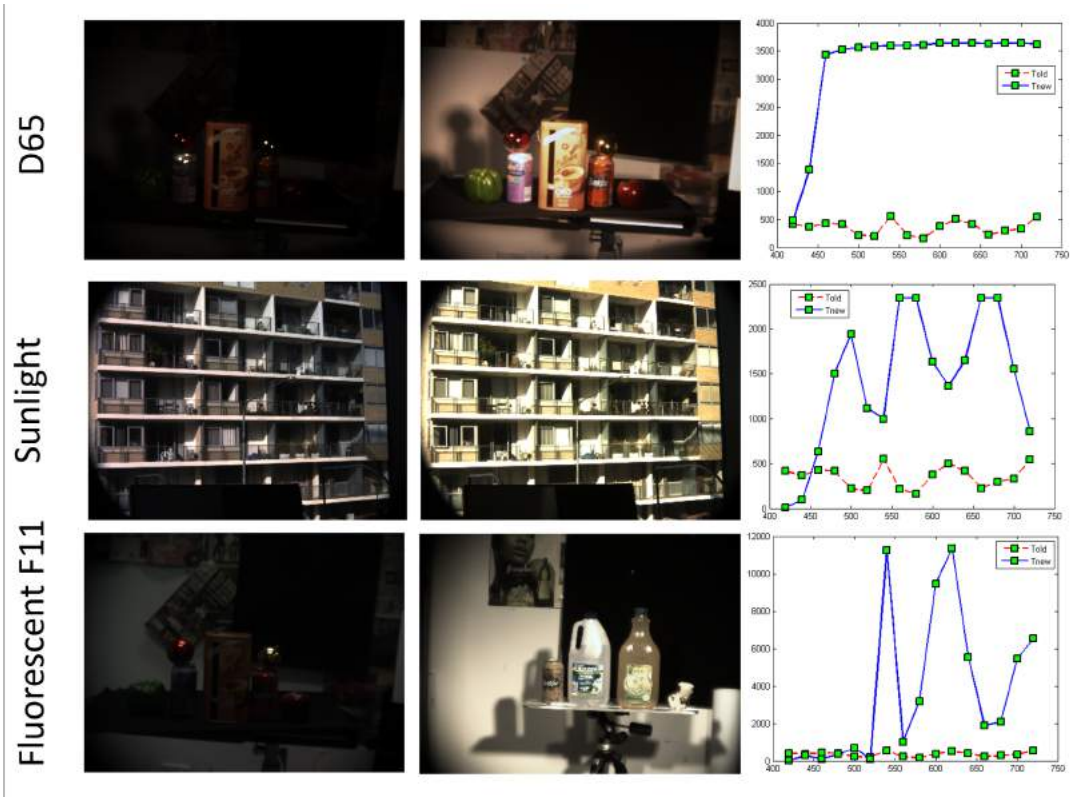

Figure 3.8: Results yielded by our method for different lighting conditions (D65, sunlight and fluorescent - F11). Left-hand column: Pseudo colour image for the initial imagery taken at random exposure times; Middle column: Pseudo colour image for the hyperspectral image acquired with the exposure times yielded by our method; Right-hand column: Time plots showing the initial and the exposure times delivered by our approach.

basic concepts governing the values yielded by a CCD at each pixel, we employ a histogram equalisation algorithm so as to obtain a closed form solution for computing the exposure times that correspond to a target cumulative distribution function of the brightness values per band. We have illustrated how our method performs under varying illumination conditions, different randomly selected initial exposure times and two target distributions, *i.e.* uniform and Gaussian. In the next chapter we will advance our method by introducing the spectral power image from which we generate the target and error images which are used for exposure time by employing cost function minimization.

Chapter 4

Controllable Automatic Exposure Adjustment for Cameras

In the previous chapter, we discussed our automatic exposure control technique for spectral cameras using histogram equalization. This technique without loss of any generality can be extended to the trichromatic cameras. In this chapter, we advance our previous approach. Here, instead of applying histogram equalization to each channel, we compute a spectral power image and apply contrast adjustment using histogram equalization, so as to calculate the exposure time for each channel using a minimization approach.

Here, we compute the exposure time based upon a formal stability analysis which assures controllability conditions are satisfied. This, in turn, yields a method which can be applied to a broad range of cameras, with varying architectures and acquisition schemes. Moreover, in order to improve the stability of the method presented here, an error image is used for the exposure time estimation. This error image is the difference between the target image and the spectral power image of the current scene, as yielded by the CIE 1931 photopic function. The method presented here has a number of advantages with respect to that previously discussed. Firstly, its controllability is assured. Secondly, the error image used here is an additional constraint as compared to the target image employed previously. Finally, the use of control theory and the error image allows for the introduction of a regularisation term in our cost function. This, in effect, implies that the optimisation problem solved here is different from that explained previously.

The chapter is organised as follows. In Section 4.1 we develop our method for automatically estimating the exposure time. In Section 4.2, we present our stability analysis for the state space representation of the method and elaborate on its controllability conditions. In Section 4.3 we show the applicability of our method for setting the exposure of a single CCD RGB camera, multiple-CCD multispectral camera and a staring arrays hyperspectral camera. In our experiments, we used a Macbeth ColorChecker to validate the quality of the results. Finally, in Section 4.4 we conclude and summarize the research presented in the chapter.

4.1 Method

Here we continue from Equation 3.2 and 3.3 given as

$$I(v,\lambda) = E(v,\lambda)Q(\lambda)T_j \ , \tag{4.1}$$

$$E(v,\lambda)Q(\lambda) = \mathcal{V}(v,\lambda) \ . \tag{4.2}$$

We can estimate the exposure time by comparing the image yielded by the current exposure time setting, with respect to an "ideal" target image. Also, recall that, the luminous flux of light sources are based upon the CIE Photopic function Stiles and Burch [1955]. CIE photopic function depicts the spectral sentsitivity response of human visual system to brightness. The CIE photopic response is presented in figure 4.1. This function is a smooth one across the spectral domain and, hence, here we also adopt the notion that the exposure times should not change steeply across adjacent wavelength-indexed bands. At the same time, we also require the method to be stable, assuring convergence. In the following subsection, we present our method to estimate the exposure time. We also present a controllability analysis.

Considering Equations (4.1) and (4.2), the initial image acquired from the spectral camera can be written as:

$$I(v,\lambda) = \mathcal{V}(v,\lambda)T_j \ . \tag{4.3}$$

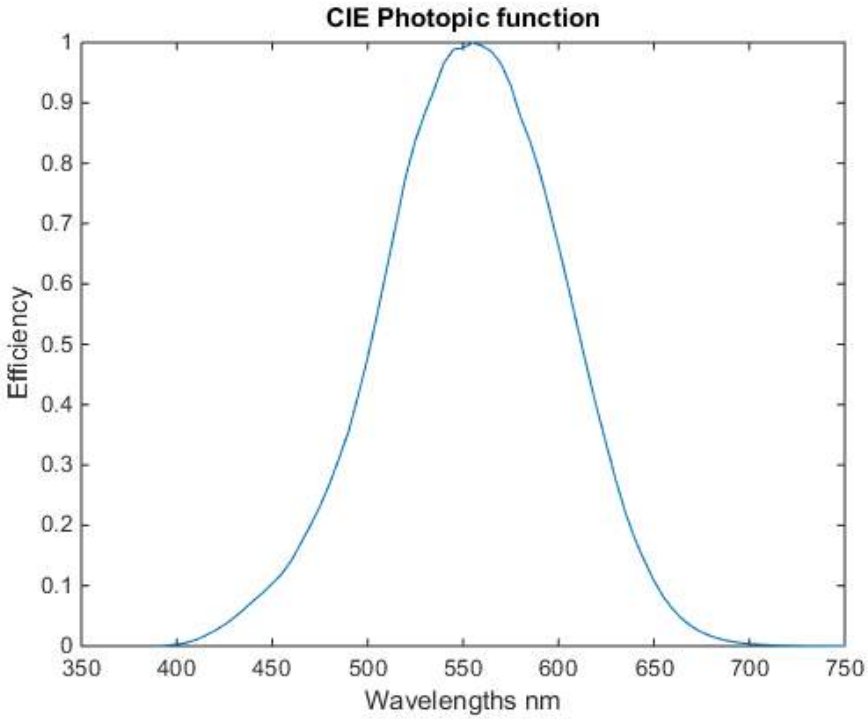

Figure 4.1: Photopic response function depicting spectral sensitivity of human visual system when exposed to brightness .

Note that spectral images can be viewed as "cubes", where the x and y axis represent spatial coordinates, and the z axis represents bands in the spectral domain. Having the initial image cube available, we can generate a target image from which subsequent updates can be computed via an optimization approach. To this end, we employ a spectral power image, which is computed making use of the CIE photopic function (Stiles and Burch [1955]) so as to weight the contribution of each spectral band to the power at each pixel. This yields:

$$\hat{I}(v) = \sum_{j=1}^{c} \sum_{\lambda \in \Lambda_j} \mathcal{V}(v, \lambda) W(\lambda) T_j \,, \tag{4.4}$$

where $W(\lambda)$ represents the photopic function at channel λ. Please note that our method is quite general in nature, allowing for other spectral sensitivity functions to be applied instead. Here, we have used the photopic function due to both its widespread use in photography and its relation to the response of the human eye

in well lit environments (Burton et al. [1992]). Similarly, in low-light conditions, a scotopic function can be used. Other choices of functions can be made based on the application domain.

It can be observed that the spectral power image $\hat{I}(v)$ is obtained by adding the weighted response for each channel. For simplicity, we can write

$$Y(v,\lambda) = \mathcal{V}(v,\lambda)W(\lambda)\ , \tag{4.5}$$

such that $\hat{I}(v)$ becomes

$$\hat{I}(v) = \sum_{j=1}^{c} \sum_{\lambda\in\Lambda_j} Y(v,\lambda)T_j\ . \tag{4.6}$$

Assuming N pixels in a single channel, Equation (4.6) can be rewritten in matrix form as follows

$$\hat{I} = \mathbf{X}T\ , \tag{4.7}$$

where $\hat{I} = [\hat{I}(1), \hat{I}(2), \ldots, \hat{I}(N)]^{\mathbf{T}}$ and $T = [T_1, T_2, \ldots, T_c]^{\mathbf{T}}$ are vectors and $\mathbf{X}$ is a matrix defined as

$$\mathbf{X} = \begin{bmatrix}\mathbf{y}_1, & \mathbf{y}_2, & \ldots & \mathbf{y}_m\end{bmatrix} \begin{bmatrix}\mathbf{1}_{\mathbf{\Lambda_1}}, & \mathbf{1}_{\mathbf{\Lambda_2}}, & \ldots & \mathbf{1}_{\mathbf{\Lambda_c}}\end{bmatrix} \tag{4.8}$$

where $\mathbf{y}_i = [Y(1,\lambda_i), Y(2,\lambda_i), \ldots, Y(N,\lambda_i)]^{\mathbf{T}}$ for $i = 1, 2, ..., m$. $\mathbf{1}_{\mathbf{\Lambda_j}}$ is an indicator function which results in a binary column vector of length m such that $\mathbf{\Lambda_j}$ is a subset of $[\lambda_1, \lambda_2, \ldots, \lambda_m]^T$ and $\mathbf{1}_{\mathbf{\Lambda_j}}(\breve{\mathbf{i}}) = 1$ if $\lambda_i \in \Lambda_j$ and 0 otherwise. Thus, $\mathbf{X}$ captures the combined effect of the irradiance making use of the photopic response per exposure time for each wavelength set Λ_j.

In order to obtain a target image for our optimization strategy, we use histogram equalization as applied to the spectral power image $\hat{I}$. We use histogram equalization because this effectively "stretches" the input image intensities across the dynamic range uniformly, resulting in a better contrasted image. It should also be noted, however, that histogram equalization produces unrealistic results for the cases where the dynamic range of the scene is significantly low. Also, the histogram-equalized image serves as a reference for our method and, therefore, other approaches can be

used as dictated by the application or lighting conditions.

Our histogram equalized image becomes

$$\mathcal{I} = histeq(\hat{I}) \ , \tag{4.9}$$

where $histeq(\cdot)$ is the image equalization operator of choice.

To estimate the update in the exposure time, we employ the difference between the input and the target image given by

$$\epsilon\left[k\right] = \mathcal{I} - \hat{I}\left[k\right] \ , \tag{4.10}$$

where $\epsilon\left[k\right]$ and $\hat{I}\left[k\right]$ are the error and spectral power images at iteration k, respectively. Note that, in the equation above, and throughout the paper, we opt to use the notation commonly found in time series analysis for the iteration indexing of the variables. We have done this for the sake of consistency with respect to the common treatment of state space analysis in the control literature.

Since the input images are all of the same scene, we can consider the target equalized images to be invariant. In practice, these may also be indexed to iteration number in a straightforward manner. For the sake of simplicity, we consider the ideal case and treat $\mathcal{I}$ as being devoid of k. It is worth noting that in many modern cameras, metering methods like centered, spot or matrix metering are used. These methods apply various weighting techniques to different parts of the scene, which can be incorporated effortlessly by the introduction of a per-pixel prior on the difference equation above.

From Equation (4.7), it can be observed that the image values are related to their exposure time as set in the matrix $\mathbf{X}$. As a result, we can use Equation (4.10) to relate the change in the exposure time to the error at iteration k as follows

$$\epsilon\left[k\right] = \mathbf{X}\Delta\left[k\right] \ , \tag{4.11}$$

where $\Delta\left[k\right]$ is the update in the exposure time. Therefore, the updated exposure time

can be calculated using the rule

$$T\left[k+1\right] = T\left[k\right] + \alpha\Delta\left[k\right] \ , \tag{4.12}$$

where α is the update step size, which suggests the magnitude of the change. Note that the larger the value of α, the steeper will be the change in the exposure time.

The exposure times are related to the camera architecture through the wavelength sets in each Λ_j. This is as these wavelengths correspond to each of the exposure times T_j, which, in turn depend on the imager. We employ a cost function that aims at both, minimizing the error $\epsilon[k]$, and also penalizing exposure times which are far removed from the average across the imager, *i.e.* we regularise T_j across all the wavelength sets. Our cost function is defined as

$$J(\mathcal{T}) = \epsilon^2[k] + \eta \sum_{j=1}^{c} \left[\gamma - \mathcal{T}_j\right]^2 \ , \tag{4.13}$$

where $\mathcal{T}_j$ is the balanced exposure time for wavelength set Λ_j, which needs to be estimated, γ is the average of exposure times such that $\gamma = \frac{1}{c}\sum_{j=1}^{c} T_j$, where $[T_1, T_2, ..., T_c]^{\mathbf{T}} = T[k+1]$ and is estimated in Eq. (4.12). η is a scalar that controls the contribution of the second term of the minimization in hand. Using Eqs. (4.7) and (4.10), the cost function can be represented in matrix form as follows

$$J(\mathcal{T}) = \left(\mathcal{I} - \mathbf{X}\mathcal{T}\right)^T \left(\mathcal{I} - \mathbf{X}\mathcal{T}\right) + \left(\sqrt{\eta}\left[\Gamma - \mathcal{T}\right]^T\right)\left(\sqrt{\eta}\left[\Gamma - \mathcal{T}\right]\right) \ . \tag{4.14}$$

Here, each column of $\mathbf{X}$ represents a wavelength-indexed channel multiplied by its photopic function value and Γ is a column vector whose entries are given by γ. Equation (4.14) can be simplified and rewritten as

$$J(T) = \begin{bmatrix} \mathcal{I} - \mathbf{X}\mathcal{T} \\ \sqrt{\eta}\Gamma - \sqrt{\eta}\mathcal{T} \end{bmatrix}^{\mathbf{T}} \begin{bmatrix} \mathcal{I} - \mathbf{X}\mathcal{T} \\ \sqrt{\eta}\Gamma - \sqrt{\eta}\mathcal{T} \end{bmatrix} \ , \tag{4.15}$$

which yields the minimization

$$\begin{aligned} \min_{\mathcal{T}} \quad & \left\| \begin{bmatrix} \mathcal{I} \\ \sqrt{\eta}\Gamma \end{bmatrix} - \begin{bmatrix} \mathbf{X} \\ \sqrt{\eta}\mathbf{I} \end{bmatrix} \mathcal{T} \right\|^2 , \\ \text{s.t.} \quad & \mathcal{T} \geq 0 \end{aligned} \tag{4.16}$$

where $\mathbf{I}$ represents the identity matrix and, since the exposure times cannot be negative, we have imposed explicitly a non-negativity constraint.

The stopping criterion for our iterative updates is governed by the normalized error given by

$$\hat{\epsilon}[k] = \frac{|\epsilon[k]|}{|\mathcal{I}|} \tag{4.17}$$

If $\hat{\epsilon}[k]$ is smaller than a predefined threshold, we can conclude convergence has been reached. This is in accordance with the notion that, if the exposure times are optimum, the difference between the spectral power image and the histogram equalized scene is expected to be nil.

4.2 Stability Analysis

Our objective as mentioned earlier is to update the exposure time such that the spectral power image is in good accordance with the histogram-equalized reference. In order to ensure that the presented method is, indeed, stable, *i.e.* that our approach will yield exposure time updates to convergence, here we present a stability analysis based upon state space. This also allows for the determination of the conditions upon which our approach is to remain controllable, *i.e.* will not diverge. Thus, in this section, we commence by transforming the equations above into their state space form Katsuhiko [2010] and then apply the stability and controllability criteria using concepts from control theory.

4.2.1 State Space Representation

State space representations are widely used in control theory so as to analyze systems with multiple inputs and outputs by making use of state variables. Recall that the

general form of a state space representation is given by the equations

$$\mathbf{x}[k+1] = \mathbf{A}\mathbf{x}[k] + \mathbf{B}\mathbf{u}[k] \ , \tag{4.18}$$

$$\mathbf{y}[k] = \mathbf{C}\mathbf{x}[k] + \mathbf{D}\mathbf{u}[k] \ . \tag{4.19}$$

Equation (4.18) is often called the state equation whereas Equation (4.19) is referred to as the output equation. In the above expressions $\mathbf{x}$ is a vector containing the state variables, $\mathbf{u}$ is the input vector, $\mathbf{A}$ is the transition matrix between state variables, the matrix $\mathbf{B}$ captures the relationship between input and state variables, the matrix $\mathbf{C}$ relates of the state variables with the outputs and $\mathbf{D}$ is the transfer matrix between the input variables and the output.

For our approach, Equations (4.10) and (4.12) are those required for the state space representation. To this end we can write

$$\begin{bmatrix} \hat{I}[k+1] \\ T[k+1] \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{I}[k] \\ T[k] \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathcal{I} \\ \epsilon[k] \end{bmatrix} \ , \tag{4.20}$$

where $\mathbf{0}$ is a matrix whose entries are all zero, $\mathbf{I}$ is the identity matrix and $\mathbf{R} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$.

The output equation is given by

$$\mathbf{y}[k] = \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{I}[k] \\ T[k] \end{bmatrix} \ . \tag{4.21}$$

We assume we have r inputs, s states, and q outputs. From above equations it can be observed that $r = \mid \mathcal{I} \mid + \mid \epsilon[k] \mid$; $s = \mid \hat{I} \mid + c$; and $q = c$, where $\mid . \mid$ yields the number of elements in the array or vector under consideration. Please note that if p number of pixels are used instead of complete image then $r = p + p$. For the reference of the reader, in Table 4.1, we summarize the variables in the state space and output equations with their respective value and dimensionality.

Equation Variables	Description	Dimensionality	Values
$\mathbf{x}$	Vector of state variables	$s \times 1$	$\begin{bmatrix} \hat{I}[k] \\ T[k] \end{bmatrix}$
$\mathbf{u}$	Vector of input values	$r \times 1$	$\begin{bmatrix} \mathcal{I} \\ \epsilon[k] \end{bmatrix}$
$\mathbf{A}$	Transition matrix between state variables	$s \times s$	$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$
$\mathbf{B}$	Relationship matrix between input and state variables	$s \times r$	$\begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$
$\mathbf{C}$	Relationship matrix between state variables and output	$q \times s$	$\begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix}$
$\mathbf{D}$	Transfer matrix between input and output	$q \times r$	$\begin{bmatrix} \mathbf{0} & \mathbf{0} \end{bmatrix}$

Table 4.1: Equivalence between the variables in the state space equations and those used in Section 4.1. Here we have $r = \mid \mathcal{I} \mid + \mid \epsilon[k] \mid$; $s = \mid \hat{I} \mid + c$; and $q = c$, where $\mid . \mid$ yields the number of elements in the array or vector under consideration.

4.2.2 Stability Analysis of the method

In control theory, a system can be either stable or not based upon its output. In state space representations, the stability of a system is determined by analyzing the eigenvalues of the matrix **A**. As mentioned earlier, the matrix **A** determines the relationship between the current state variables and those at the next iteration. A discrete time system will be unstable if any of the eigenvalue of matrix **A** has a magnitude greater than unity. On the other hand, a discrete time system will be stable if all the eigenvalues of matrix **A** have magnitudes less than 1. For a marginally stable system, the eigenvalue magnitudes are equal to 1. Those eigenvalues whose magnitudes are unity, must also be simple roots of the characteristic polynomial of

A. In other words, the Jordan blocks of matrix **A** must be of first order (Katsuhiko [2010]) .

From Table 4.1, it becomes evident that the eigenvalues of matrix **A** are either 0 or 1. It can also be observed that unity eigenvalues correspond to the simple root of characteristic polynomial of **A**. Furthermore, it can be shown that the number of nil eigenvalues relates to the number of pixels in $\hat{I}$, whereas the number of unity eigenvalues is determined by the number of exposure times used by the system.

Thus, if, for instance, the system is a multiple CCD one, the number of unity eigenvalues is equal to the number of CCDs. If the system exposure time is channel-based, *i.e.* one exposure time per wavelength indexed band, then the amount of unity eigenvalues equals the number of bands in the image. Further, these unity eigenvalues suggest that our method is marginally stable (the state of the system will converge with "oscillations" to the optimal exposure time when the iteration number tends to infinity). Note that this is not a problem in practice due to the use of the stopping criterion presented earlier. This is the case since the threshold allows for the method to stop even when these oscillations may be present in its asymptotic behavior.

4.2.3 Controllability

Another important aspect of the system which is worth exploring is the controllability of the method. Controllability helps to determine the effect of the inputs on the system. A state is controllable at iteration k if there exists an input that delivers the desired output as $k \to \infty$.

In order to determine the controllability of the system, a controllability matrix **P** is usually constructed using the matrix **A** and the input matrix **B**. The matrix **P** is given by

$$\mathbf{P} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{s-1}\mathbf{B} \end{bmatrix} , \tag{4.22}$$

where the dimension of $\mathbf{P} = s \times rs$.

For all the states to be controllable, matrix **P** is required to be full row rank. This refers to the case whereby all the states will converge to the desired state asymptotically. If the matrix **P** is found to be rank deficient, then such states need to be

separated from the rest and analyzed in order to determine those inputs which render them uncontrollable.

The matrix $\mathbf{P}$ can be found by first estimating the state space matrix products, $\mathbf{AB}$, $\mathbf{A}^2\mathbf{B}$, ..., $\mathbf{A}^{s-1}\mathbf{B}$. For our method, these are given by

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \\
\mathbf{A}^2\mathbf{B} &= \mathbf{A}(\mathbf{AB}) \\
&= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \\
&\vdots \\
&\vdots \\
\mathbf{A}^{s-1}\mathbf{B} &= \mathbf{A}(\mathbf{A}^{s-2}\mathbf{B}) \\
&= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}
\end{aligned}$$

Using the matrices above, the controllability matrix $\mathbf{P}$ becomes

$$P = \begin{bmatrix} \mathbf{I} & \mathbf{I} & |\mathbf{0} & \mathbf{0}| & \dots |\mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & |\mathbf{0} & \mathbf{R}| & \dots |\mathbf{0} & \mathbf{R} \end{bmatrix} \tag{4.23}$$

Note that the matrix $\mathbf{P}$ depends on the matrix $\mathbf{R}$ in terms of its row rank. Since $\mathbf{R} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, it can be observed that $\mathbf{R}$ is derived from $\mathbf{X}$, which contains the irradiance weighted by the photopic response for each exposure time. Moreover, for p pixels and c channels, $\mathbf{X}$ has dimensions $p \times c$ and $\mathbf{R}$ has dimensionality of $c \times p$. Thus, in order for our method to remain controllable, the matrix $\mathbf{R}$ should be full row rank. As a result, our method requires at least c linearly independent pixels over the whole range of wavelength-indexed bands for Equations (4.10) and (4.12) to remain controllable. This is due to the fact that using at least c linearly independent pixels across the c channels under consideration implies the matrix $\mathbf{R}$ has c linearly independent rows. This also assures the column rank of $\mathbf{P}$ is at least $c + p$, which in

turn satisfies the controllability condition for our method.

Hence, typical uncontrollable cases would be those in which, for instance, a single pixel is used. Recall that, for a single pixel, **X** becomes a vector and, as a result, **P** becomes row rank deficient. Other cases are those where the matrix **P** is row rank deficient due to linear dependency between the pixels. This is the case, for instance, when every channel or CCD captures a completely dark image, which corresponds to $\mathbf{X} \equiv 0$, and thus every row of **R** is also nil.

4.3 Results and Discussion

In Algorithm 1, we show the step sequence for our algorithm. Note that the algorithm departs from an initial exposure time and iterates until convergence is reached (the normalized error is below a pre-defined threshold). Moreover, it is worth noting in passing that, in practice, these iterations are not overly computationally intensive. The worst case bound of the algorithm is $O(N^3)$, assuming the input image is of spatial dimensions NXN. This is since the problem is essentially a constrained least squares optimization one, which has been thoroughly studied in the numerical methods literature (Stark and Parker [1995]). For Step 6, we have applied the histogram equalization method of Bassiou and Kotropoulos [2007] with a uniform target histogram. We would like to stress that the choice of method to compute the reference image in Step 6 or equalization technique is not prohibitive and other methods elsewhere in the literature may be also used without any loss of generality.

We have tested our method using two different imagers. The first of these is a Fluxdata 3CCD multi-spectral camera. The second imager is based upon an liquid crystal tunable filter and a QImaging Retiga camera. This system acquires 33 bands in the range between $380nm$ and $700nm$ in intervals of $10nm$. Note that these cameras vary significantly in their architecture. This is as the former is a multispectral imager where multiple channels share the same exposure time, which is determined by each of the three CCDs comprising the camera. The latter is a staring array system where every wavelength channel can be assigned an individual exposure time. For our experiments, we have placed a Macbeth ColorChecker in all the scenes under study.

Algorithm 1 Estimate Exposure Time

procedure EXPOSURE(T_i)
  Initialize the exposure time when $k = 1$, $T[k] \leftarrow T_i$
  for $k = 1,...,maxIterations$ **do**
    Acquire Cube image $I(v, \lambda)$
    Generate Spectral Power Image Using Photopic function in Eq. 4.4 $\hat{I}(v) \leftarrow I(v, \lambda)$
    Generate target image using histogram equalization in Eq.4.9 : $\mathcal{I} \leftarrow \hat{I}$
    Compute $\Delta[k]$ using Eq.4.11
    Find $T[k+1]$ using $\Delta[k]$ using Eq. 4.12
    Balance $T[k+1]$ using cost function in Eq.4.16 : $T[k+1] \leftarrow \mathcal{T}$
    Find normalized error $\hat{e}$ using Eq.4.17
    if $\hat{e} \leq threshold$ **then**
      break for
    end if
  end for
  Return $T_{final} \leftarrow T[k+1]$
end procedure

These color checkers are widely used in professional photography to perform white balancing or color calibration (McCamy et al. [1976]). For purposes of exposure time setting, we use the gray-scale tiles at the bottom row of the chart, which contain six gray-scale colors. These vary from white to black in regular intensity intervals. A well contrasted image can be acquired by setting the exposure such that the photopic response with respect to these gray scale tiles is linear while maximizing the dynamic range of the camera. In all our experiments we have set the initial exposure times to $50ms$, threshold to 0.3 and the update step size $\alpha = 0.1$

In Figure 4.2, we show the results yielded by our method when acquiring two scenes using the 3-CCD multispectral camera. For each scene we show the spectral power image $\hat{I} = \mathbf{X}T$ of the scene at the initial, middle and final iterations of our method. Next to each spectral power image, we show its photopic response plot for the gray scale tiles of the Macbeth ColorChecker. In the plot, the x-axis represents the index of gray scale tiles, *i.e.* from black to white, indexed 1 to 6. The y-axis represents the photopic response. At the bottom of each scene, we show the normalized error as a function of iteration number.

From the figure, we can note that the initial spectral power images are noticeably

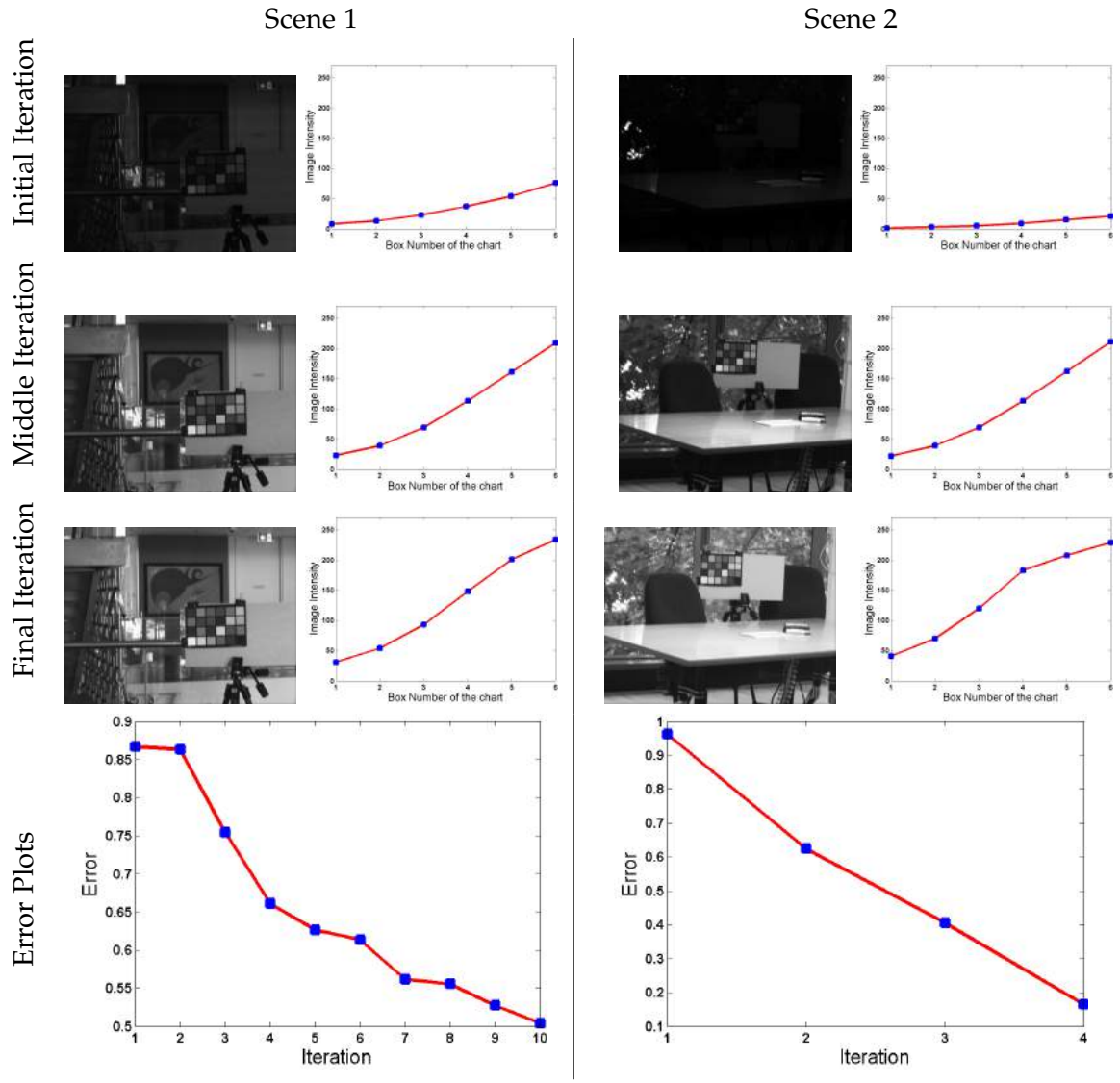

Figure 4.2: FluxData multi-spectral camera scenes. Left-hand column: Spectral power image; Right-hand column: Color chart gray-tile intensity plots (the x-axis corresponds to the index of the tile and the y-axis to the photopic response). In the figure, the top row shows the initial images, whereas the second row corresponds to the images produced by the exposure times at the iteration corresponding to the middle point over the convergence of our method. The third row shows the image acquired using the exposure time yielded by our method after convergence has been reached. Finally, the bottom row shows the normalised error plots as a function of iteration number.

underexposed and, as our method converges, they gradually improve. This is also captured by their photopic response function values as plotted in the figure. Further, as our method iterates to convergence, the contrast of the spectral power image improves, with a photopic response over the gray tiles on the color chart approximating a linear response over the whole dynamic range of the camera. Also, from the error plots, we can observe that the method starts with a high error value, which gradually decreases towards the threshold. Note that, the integration times of spectral cameras are higher than traditional cameras, due to larger number of spectral channels. Therefore, for our method, the initial exposure time has been chosen to be small by default, so as to minimize the integration time. Whereas, the method is equally capable of using overexposed images to reach to the target setting.

In Figure 4.3 we repeat this sequence for two real-world scenes acquired using the LCTF-based imager. Consistent with our results shown in Figure 4.2, our method steadily converges towards the threshold error departing from a typically high value. Also, note that the images delivered at output depict a much improved contrast as compared to those taken with the initial exposure times. Moreover, the exposure times delivered by our method are almost linear over the dynamic range of the CCDs in the camera.

In Figure 4.4 we show in first column the baseline images acquired by manually setting the exposure times such that the photopic response over the gray tiles (shown in second column) approximate a linear function over the whole dynamic range of the camera. Please note that in some cases the scene lighting conditions in conjunction with the camera setup do not permit the plots to be completely linear. As a result, our baseline images were manually exposed so as to obtain the best linear relationship between the gray tile brightness and the CCD dynamic range. In Figure 4.4, we show the absolute error pixel map of our method at the beginning, middle, and end of the iterative process in Algorithm 1. The first two columns correspond to the scenes acquired by the FluxData multi-spectral camera as shown in Figure 4.2. The last two columns correspond to the two scenes acquired by the OKSI hyperspectral camera shown in Figure 4.3. From the figure, it can be observed that the error per pixel also reduces as the method iterates.

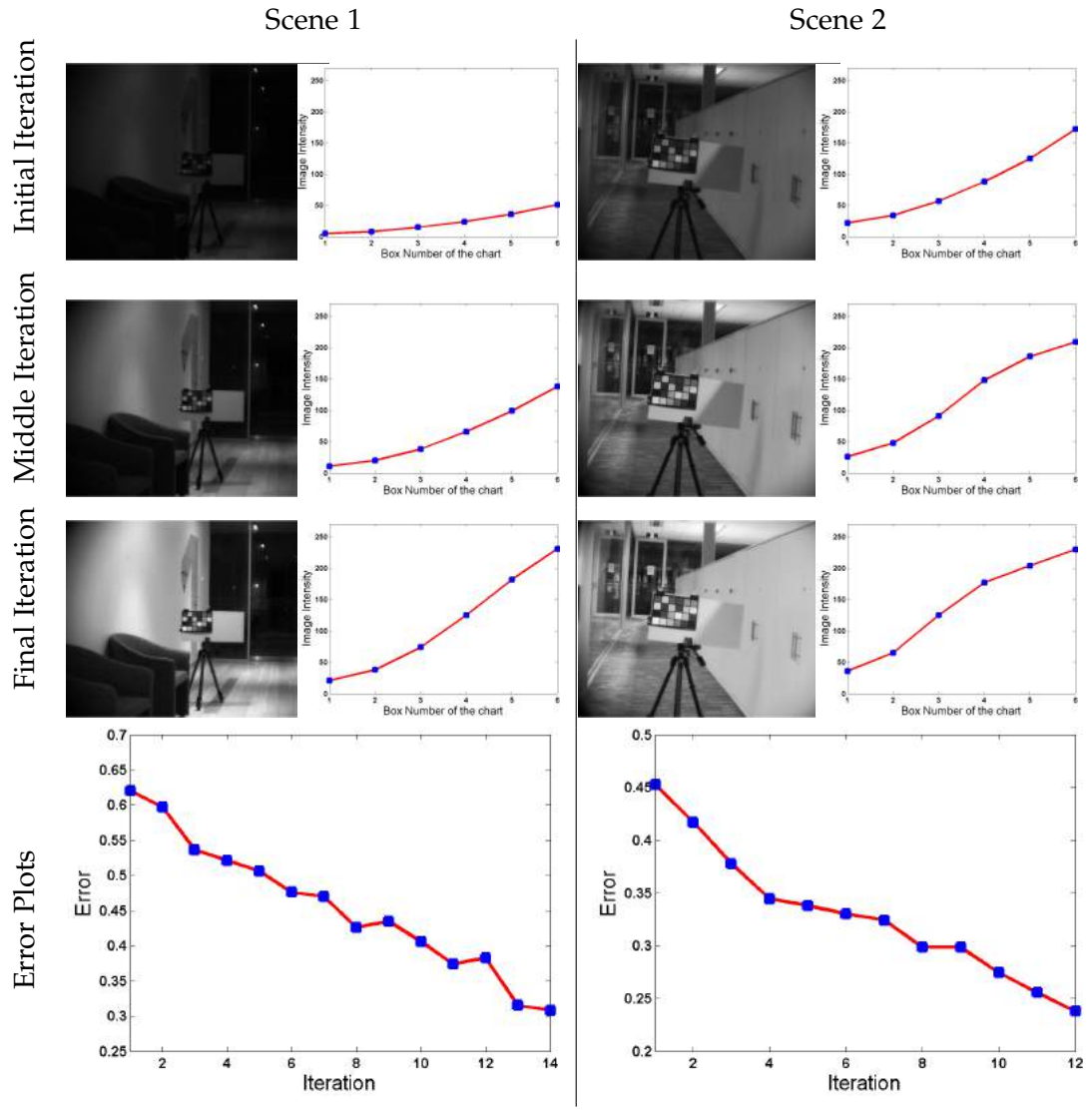

Figure 4.3: OKSI hyperspectral camera scenes. Left-hand column: Spectral power image; Right-hand column: Color chart gray-tile intensities (the x-axis corresponds to the index of the tile and the y-axis to the photopic response). In the figure, the top row shows the initial images, whereas the second row corresponds to the images produced by the exposure times at the iteration corresponding to the middle point over the convergence of our method. The third row shows the image acquired using the exposure time yielded by our method after convergence has been reached. Finally, the bottom row shows the normalised error plots as a function of iteration number.

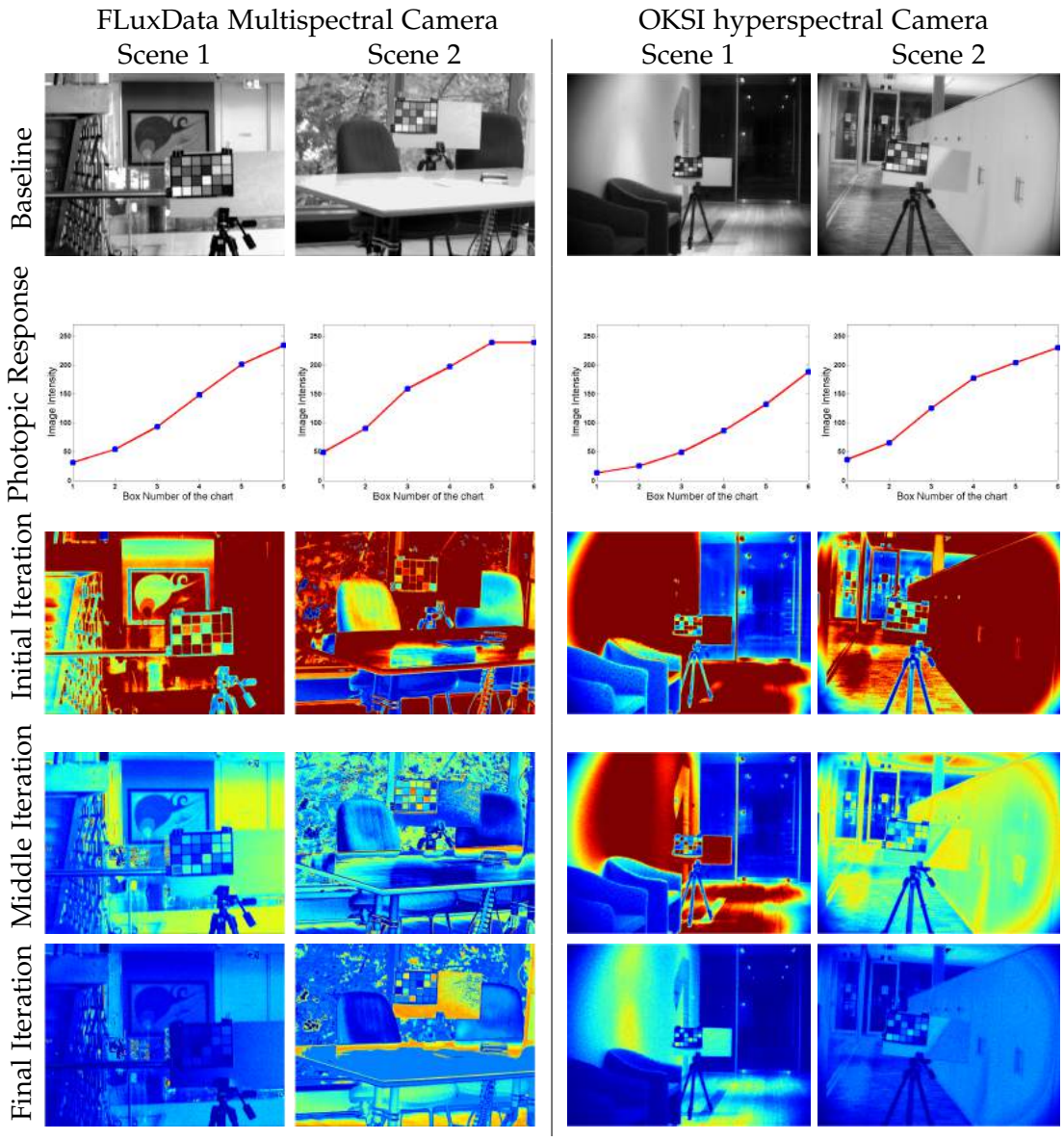

Figure 4.4: Baseline images acquired by setting exposure times manually with their photopic response of gray tiles, and per pixel error map of our method at different iterations. From top-to-bottom: Baseline image, plot of photopic response of gray tiles in the baseline image, per pixel error map of images at initial, middle and end iterations of our method. Left-hand columns : Scenes acquired by the FluxData multi-spectral camera; Right-hand columns: scenes captured using the OKSI hyper-spectral camera. All the error maps are normalized such that blue corresponds to zero and red to unity.

Figure 4.5: Spectral power images for the color chart using the exposure times obtained manually (left) and those obtained using our method (right).

For testing our method on RGB we take the output of a single CCD on Fluxdata camera which has a bayer filter to generate an RGB image. For purposes of exposure time setting, we use the gray-scale tiles at the bottom row of the chart, which contain six gray-scale colors.These vary from white to black in regular chromaticity intervals. Thus, for our baseline imagery, we manually adjust the exposure of the camera so as to achieve a linear photopic response with respect to these grayscale tiles while maximizing the dynamic range of the camera, *i.e.* a linear response between 0 to 255. The essence lies in CIE photopic response, which estimates the average spectral sensitivity of human visual system, for which a well exposed image represents all the dark and bright regions equally, with a smooth transition between shades. Thus we take the baseline image as the one which has a linear and stretched response over the gray tiles of spectral power image.

In Figure 4.5 we show results using the Macbeth ColorChecker illuminated by a calibrated incandescent illuminant. In the figure we show the spectral power image of the input RGB image for the baseline color chart image (left) and that obtained using the exposure times yielded by our method (right). We also provide quantitative analysis in Figure 4.6, where we show the plot of the photopic response as a function of the indices of gray-scale tiles on the chart ordered in increasing brightness for both, the manually exposed image (broken line) and that taken using the exposure times yielded by our method (solid line).

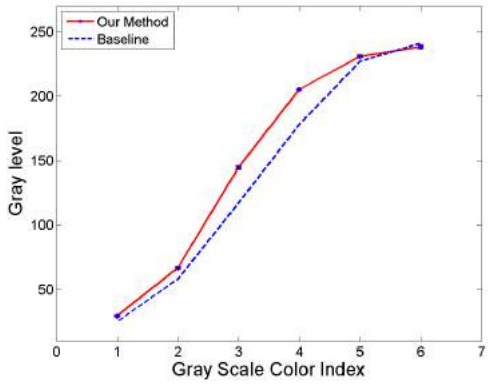

Figure 4.6: Photopic response as a function of the indices of the gray-scale tiles in the Macbeth ColorChecker for our method (solid line) and the manually exposed image (broken line).

Note that, in the plots, the photopic response yielded by our method is in good accordance with that corresponding to the manually exposed image. Moreover, the plot covers an interval which is in close accordance with the dynamic range of the camera. Comparing the two plots, the main difference stems from a slight increase in the exposure times obtained by our method for the gray levels on the center-section of the chart. This difference can also be seen in the third and fourth tiles of the bottom row of the images in Figure 4.5.

Finally we show results of RGB images on real-world imagery. To this end, we have captured images under different lighting conditions in both indoor and outdoor environments. As before, we have captured two sets of imagery. The first of these corresponds to the images acquired when the ColorChecker is used so as to manually obtain a linear photopic response. The second set is that corresponding to the same scene captured using the exposure times recovered by our method.

In Figure 4.7 left-most column, we show the spectral power image of a scene when color chart is placed and its exposure time is adjusted manually such that a stretched and linear photopic response is obtained from the gray tiles. Using the manually adjusted exposure time a base line image is captured which is shown in the middle column of the figure. In the rightmost column we show the RGB images obtained for the same scenes with exposure time setting calculated by our method.

Note that the two sets of images (i.e the baseline and our method resultant) are barely distinguishable from each other. This implies that the timings recovered by our method and those calculated manually using the Macbeth chart are in close accordance.

4.4 Summary

In this chapter we have presented a method for the automatic computation of the exposure times for spectral cameras. Here, we have shown how the exposure time can be recovered by using a regularised non-negative least squares optimization. We have also presented a controllability and stability analysis using the state space representation of our method and illustrated the applicability of our approach for purposes of automatic exposure control of staring array and multi-CCD and single CCD cameras. In the next chapter we will introduce our proposed technique for camera radiometric calibration. This radiometric calibration is important in order to have accurate radiance representation of the scene. Our radiometric calibration method can be applied to cameras using multiple images as well as a single well exposed image.

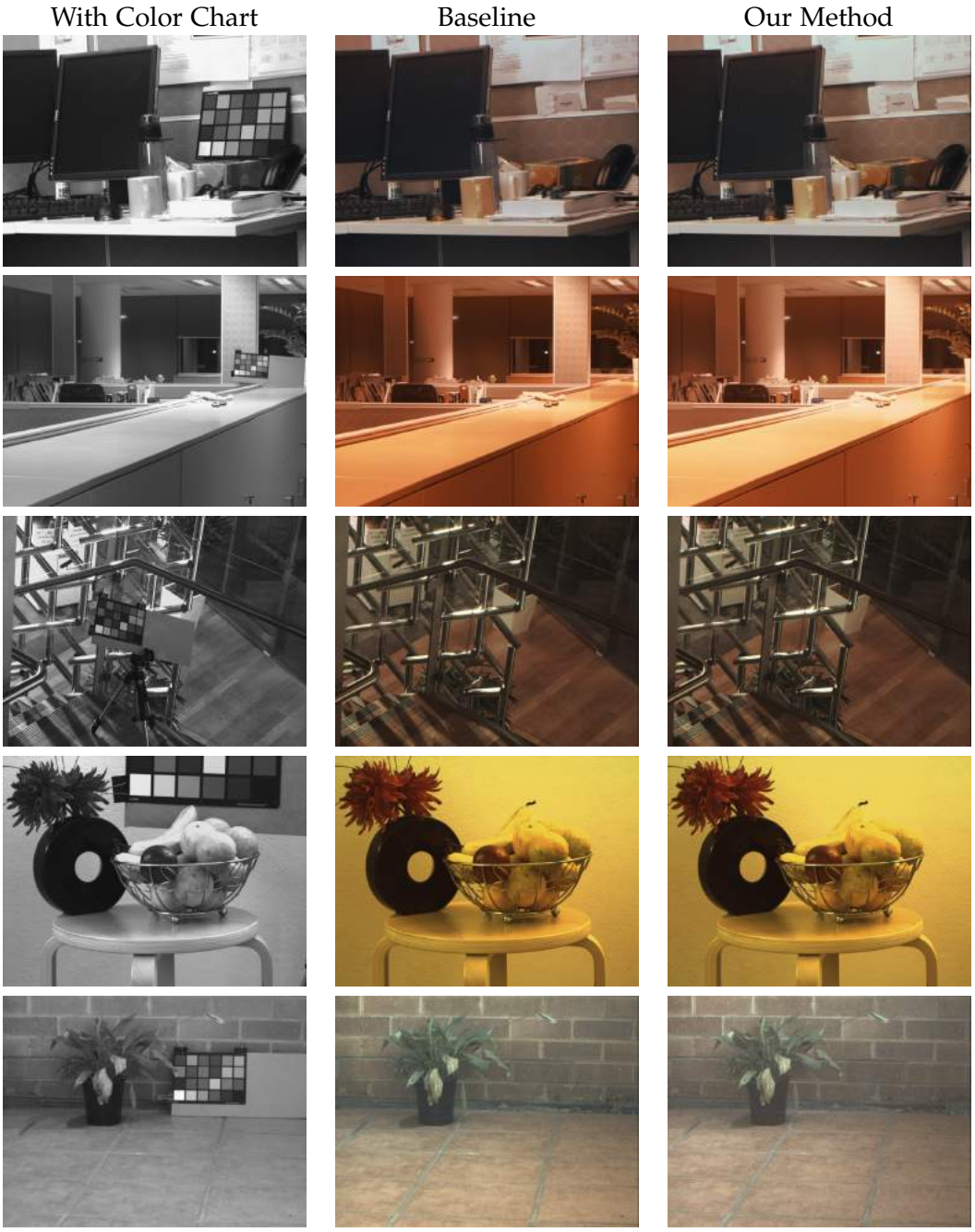

Figure 4.7: RGB images for indoor and outdoor scenes acquired by first placing color chart to determine exposure time with help of grayscale tiles (Shown in left-most column). Baseline images were acquired using manually computed exposure times shown in middle column. In the right-most column results shown are yielded by our method .

Chapter 5

Radiometric calibration

In this chapter, we propose a method to perform radiometric calibration for cameras. Since radiometric calibration is performed on each channel independently, our proposed method can be applied without loss of generality to spectral as well as trichromatic cameras. Radiometric calibration is important to translate the obtained image intensity into image radiance, which is highly desired in spectral imaging. High dynamic range imaging is a widely used technique which is generally obtained using radiance maps generated by radiometric calibration. Note that the quality of a HDR image is highly dependent on the accurate estimate of radiometric calibration. For this reason, we show the benefit and accuracy of our radiometric calibration method using trichromatic HDR images. In this chapter, our choice to show our results in RGB channels instead of a spectral cube is due to the ease of visualization of RGB images.

Thus, we view the camera response function as the geometric median with respect to the scene radiance. This allows the recovery of the camera response function and image radiance by minimising a cost function based upon a weighted L2-norm which was solved using the Lq Weiszfeld algorithm in (Aftab et al. [2015]). The Lq Weiszfeld algorithm is an iterative algorithm which uses weights to find a special point, it call x in n dimensional space that has the minimum sum of difference from the available data points. The Weiszfeld algorithm solves for the L_1 mean or geometric median of a set of points $\{y_1, y_2, ..., y_k\}$ in n dimensional space by minimizing the following

cost function

$$C_1(x) = \sum_{i=1}^{k} ||x - y_i|| \,, \tag{5.1}$$

where $||.||$ is the Euclidean norm. The Weiszfeld algorithm updates a current estimate x^t to

$$x^{t+1} = \frac{\sum_{i=1}^{k} w_i^t y_i}{\sum_{i=1}^{k} w_i^t}, \tag{5.2}$$

such that $w_i^t = ||x^t - y_i||^{-1}$. The advantage of this approach stems from the fact that this treatment can be shown to yield the solution to an equivalent L1-norm cost function which is robust to outliers. This allows us to use multiple images for radiometric calibration so as to achieve higher accuracy with lesser number of input LDR images. Our method robustly deliver HDR imagery with as few LDR images as possible. We also extend this cost function setting to use on a single LDR image for radiometric calibration. For single image radiometric calibration (SIRC), the camera response function and irradiance image are estimated. The irradiance image is acquired by applying inverse camera response function on the input LDR image. It should be noted that the radiance image is obtained from multiple image radiometric calibration (MIRC), whereas, we obtain irradiance image in SIRC. The irradiance image has the effect of exposure incorporated in it where as radiance image is free from the effect exposure. Since the main contributors to the nonlinearity comes after the imaging plane where irradiance is observed, the calibration involving the irradiance image still leads to correct camera response function. We present the approach where we iteratively estimate the camera response function and the image irradiance. In SIRC instead of multiple images as in case of MIRC we use color channel intensities whose irradiance values are closer within a range to perform iterative calibration.

In the Section 5.1, we explain our approach for multiple image that depicts the generalized representation of the method. Section 5.2 describes the initialization of the method. Section 5.3 presents single image radiometric calibration as a special case

of our method explained in Section 5.1. In Section 5.4, we discuss the experiments conducted using our proposed method and present comparison with other widely used methods. Finally Section 5.6 concludes our work on radiometric calibration.

5.1 Estimating the Camera Response and Radiance Map

To commence, we depart from the image acquisition process. If the bias and the dark current, *i.e.* the constant response of the CCD present even if no light is impinging the sensor, are neglected, we can view the scene radiance $E(u)$ at pixel u as being proportional to the photon flux $P(u)$ incident on the CCD (Robles-Kelly and Huynh [2013]), thus we can write

$$E(u) = \tau P(u) \tag{5.3}$$

where τ is a proportionality constant.

This is important since we can view the value of the u^{th} image pixel in the image as a non-linear function of the scene radiance, *i.e.* the camera response function, such that

$$I(u) = g\big(E(u)T\big) = g\left(\frac{\zeta}{\gamma}P(u)Q\alpha T\right) \tag{5.4}$$

where Q is the quantum efficiency, γ is the CCD gain, α is the camera aperture and T is the exposure time.

Consider a set of m LDR images $\mathcal{I} = \{I_1, I_2, \ldots, I_m\}$ acquired at different exposure times $\mathcal{T} = \{T_1, T_2, \ldots, T_m\}$. It is worth noting that, for the images in $\mathcal{I}$, the quantum efficiency can be assumed to remain constant. Also, due to the proportionality constant ζ in Equation 6.2, the radiance estimated here is expected to be up-to-scale with respect to the true radiance. Thus, from now on, and without any loss of generality, we assume the image radiance is normalised to unity.

As a matter of convenience, in the following equations we use the shorthand

$$f_c[I_i(u)] = E(u)T_i \tag{5.5}$$

where $f_c[I_i(u)]$ corresponds to $g^{-1}(E(u)T_i)$ for the colour channel indexed c when

the exposure time accounts for yielding the exposure corresponding to the i^{th} LDR image in $\mathcal{I}$.

Note that, by writing the problem as the minimisation

$$\min_{f_c[\cdot],E(\cdot)} \left\{ \sum_{I_i \in \mathcal{I}} \sum_{u \in I_i} (f_c[I_i(u)] - E(u)T_i)^2 \right\} \tag{5.6}$$

we can view the camera response function as the geometric median with respect to the scene radiance. This observation is important since it hints at a Weiszfeld minimisation.

Recall that Weiszfeld's algorithm is an iterative one. Moreover, it is based on a fixed point iteration. Thus, we recover the image radiance and camera response functions making use of the cost functions

$$G_c[E^{n+1}(u)] = \sum_{i=1}^{m} \left| \frac{f_c^n[I_i(u)]}{T_i} - E^{n+1}(u) \right| \tag{5.7}$$

and

$$H_c\left[f_c^{n+1}[b]\right] = \sum_{u \in \Gamma_b} \left| E^n(u)T_i - f_c^{n+1}[I_i(u)] \right| \tag{5.8}$$

where n is the iteration number and $c = \{R, G, B\}$ is the colour channel index, *i.e.* red, green or blue. In Equation 5.8, Γ_b is the set of pixels whose colour channel indexed c have the intensity value b.

To minimize the cost functions in Equations 5.7 and 5.8, we employ the Lq solution to Weiszfeld's algorithm presented by Aftab et al. [2015]. Thus, we can rewrite the cost functions in Equations 5.7 and 5.8 as follows

$$G_c^*[E^{n+1}(u)] = \sum_{i=1}^{m} \alpha_i^n \left(\frac{f_c^n[I_i(u)]}{T_i} - E^{n+1}(u) \right)^2 \tag{5.9}$$

$$H_c^*\left[f_i^{n+1}[b]\right] = \sum_{u \in \Gamma_b} \eta^n(u) \left(E^n(u)T_i - f_c^{n+1}[I_i(u)] \right)^2 \tag{5.10}$$

where α_i^n and $\eta^n(u)$ are the corresponding weights for the equivalent L2-norm cost functions.

The advantage of this treatment resides in the fact that we can now solve the cost functions in Equations 5.9 and 5.10 via a least squares minimisation with the following update rules

$$E^{n+1}(u) = \frac{\sum_{i=1}^{n} \alpha_i^n \frac{f_c^n[I_i(u)]}{T_i}}{\sum_{i=1}^{n} \alpha_i^n} \tag{5.11}$$

$$f_c^{n+1}[b] = \frac{\sum_{u \in \Gamma_b} \eta^n(u) E^n(u) T_i}{\sum_{u \in \Gamma_b} \eta^n(u)} \tag{5.12}$$

Following Aftab et al. [2015], the weights are updated making use of the expressions

$$\alpha_i^n = \left(\frac{f_c^n[I_i(u)]}{T_i} - E^n(u) \right)^{-1}; \tag{5.13}$$

$$\eta^n(u) = \left(E^n(u) T_i - f_c^n[I_i(u)] \right)^{-1} \tag{5.14}$$

Since these weights when applied iteratively converges to an L1 solution, the resultant camera response function and the radiance map images are robust to outliers thus reducing the error.

5.2 Initialization

Note that, since the updates above rely on the estimates yielded by previous iterations, we require an initial estimate of the radiance image and the camera response function. To this end, we commence by computing an initial estimate of the scene radiance making use of a linear camera response.

With the initial radiance map in hand, the estimation of the initial camera response function is a straightforward matter by minimizing the cost function given by

$$K_c\left[f_c^0[b]\right] = \sum_{u \in \Gamma_b} \left(E^0(u) T_i - f_c^0[I_i(u)] \right)^2 \tag{5.15}$$

where $E^0(u)$ refers to the initial radiance estimate.

Note that this is equivalent to the cost function in Equation 5.8 where the weight

$\eta^0(u) \equiv 1$. This minimization yields

$$f_c^0(b) = \frac{1}{\mid \Gamma_b \mid} \sum_{u \in \Gamma_b} E^0(u) T_i \tag{5.16}$$

which can be viewed as the mean, *i.e.* expected value, of the radiance values for all those pixels having intensities equal to the reference value b.

5.3 Single Image Radiometric Calibration

We present single image radiometric calibration (SIRC) as a special case of multiple image radiometric calibration (MIRC). We present the approach to estimate the camera response curve using a single well exposed image. We assume I to be single well exposed image, representing both dark and bright regions of the scene. Since we are dealing with a single image, the exposure time here acts as a constant and can be absorbed with normalized values, thus Equation 5.5 can be represented as :

$$f[I_i(u)] = R_i(u) \ . \tag{5.17}$$

Here, we shift our attention from a case of m multiple exposed images $\mathcal{I} = \{I_1, I_2, \ldots, I_m\}$ (MIRC), to a single well exposed image (SIRC) with multiple channels $\{r, g, b\}$. Unlike MIRC, here we assume that the CRF of all channels is same, as the difference is often negligibly small. In order to estimate the radiance to intensity mapping in MIRC, we looked for different intensities across m images for a given point, assuming they shared the same radiance value. In case of SIRC, we have a single image, thus, we look across the channels of the image, for a given point that have same radiance values for each channel. We assume that for a given point in an image, if its red, green, and blue channel intensities are close to each other, within a threshold range τ, this suggests that they have received a similar level of radiance. It is based on the assumption that, the normalized spectral sensitivity functions of camera are used (Jiang et al. [2013]), which provide a similar response by the spectral sensitivity functions of RGB channels, when exposed to similar level of radiance.

Thus in Equation 5.17 the subscript i represents color channels i.e. red, green

and blue. We have dropped subscript c from the camera response function as we assume that the CRF is approximately same for each red, green and blue channels. Here f is the camera response function to be estimated and $R_i(u)$ represents the normalized irradiance at pixel u of channel i. We view R as an irradiance image as it is normalized thus the exposure time is assumed to be absorbed in it. Similar to MIRC, here we also commence with a linear camera response function, with which initial irradiance image is calculated. Similarly, we use Weiszfeld algorithm to iteratively perform calibration. In each iteration we segment those pixels of irradiance image whose red, green, and blue irradiance values are closer to each other within the threshold range τ. We call the irradiance image which contains only these selected pixels as intermediate irradiance image defined as $\widetilde{R}$. Similarly, we use $\widetilde{I}$ for the input LDR image which contains only those pixels selected from the irradiance image whose red, green , and blue intensities are with in τ.

Here, we recover the intermediate irradiance image and camera response functions making use of the following cost functions

$$\widetilde{G}[\widetilde{R}^{n+1}(u)] = \sum_{i=1}^{3} \left| f^n[\widetilde{I}_i(u)] - \widetilde{R}^{n+1}(u) \right| \tag{5.18}$$

and

$$\widetilde{H}\left[f^{n+1}[b]\right] = \sum_{u \in \Gamma_b} \left| \widetilde{R}^n(u) - f^{n+1}[\widetilde{I}_i(u)] \right| \tag{5.19}$$

where, n is the iteration number. In Equation 5.18, i values $(1, 2, 3)$ represents (red,green,blue) respectively. In Equation 5.19, Γ_b is the set containing red, green , and blue pixels intensities equal to b or within the τ range of b. Here, $\widetilde{R}$ represents the intermediate irradiance image, which is a single channel image used for the calculation of the camera response curve.

As done earlier, we employ Weiszfeld algorithm to minimize the above cost functions which yields 5.18 and 5.19 as follows

$$\widetilde{G}^*[\widetilde{R}^{n+1}(u)] = \sum_{i=1}^{3} \alpha_i^n \left(f^n[\widetilde{I}_i(u)] - \widetilde{R}^{n+1}(u) \right)^2 \tag{5.20}$$

$$\widetilde{H}^{*}\left[f_i^{n+1}[b]\right] = \sum_{u\in\Gamma_b} \eta^n(u)\left(\widetilde{R}^n(u) - f_c^{n+1}[I_i(u)]\right)^2 \tag{5.21}$$

where α_i^n and $\eta^n(u)$ are the corresponding weights for the equivalent L2-norm cost functions updated making use of the expressions

$$\alpha_i^n = \left(f^n[\widetilde{I}_i(u)] - R_i^n(u)\right)^{-1} \tag{5.22}$$

$$\eta^n(u) = \left(\widetilde{R}^n(u) - f^n[\widetilde{I}_i(u)]\right)^{-1} \tag{5.23}$$

Using the pixels selected from the irradiance image, we update the intermediate irradiance image by minimizing Equation 5.20 with weight found in Equation 5.22 given as.

$$\widetilde{R}^{n+1}(u) = \frac{\sum_{i=1}^{n} \alpha_i^n f^n[\widetilde{I}_i(u)]}{\sum_{i=1}^{n} \alpha_i^n} \tag{5.24}$$

This intermediate irradiance image is a single channel image which we use to calculate camera response function by minimizing Equation 5.21 given as.

$$f^{n+1}[b] = \frac{\sum_{u\in\Gamma_b} \eta^n(u)\widetilde{R}^n(u)}{\sum_{u\in\Gamma_b} \eta^n(u)} \tag{5.25}$$

Since we are using a threshold range τ for adding those pixel location in our calculation whose irradiance values are closer to each other, this may add noise and outliers to our camera response function curve estimation. In order to mitigate the noise and the effect of outliers we use the curve smoothing function proposed by Cleveland [1979]. This results in points smoothed to form a camera response function curve. It is pertinent to mention that the pixels selected from the irradiance image may not be sufficient enough to represent the complete dynamic range. In order to represent the complete camera response curve we estimate the coefficients of polynomial curve that fits the pixel data obtained after smoothing, using least squares approach as mentioned in (Weisstein [2002]). This yields a smooth camera response curve which we use to estimate the irradiance image using Equation 5.17. We use this irradiance image to start the next iteration where we again select pixels

using the threshold range τ to form intermediate irradiance image and continue the above mentioned steps until the error $\sum_{u\in\Gamma_b}\left|\widetilde{R}^n(u) - f^{n+1}[\widetilde{I}_i(u)]\right|$ is minimized.

5.4 Experiments

To this end, we commence by using hyperspectral images data as radiance, so as to generate LDR images using ground truth camera response functions. This allows for both, a qualitative and quantitative analysis making use of ground truth radiance maps. Later on, we use real-world LDR images to generate HDR panoramas.

For our quantitative analysis, we have used three hyperspectral image datasets. The first of these is that presented in by Foster et al. [2006]. This dataset contains 8 outdoor reflectance scenes. The other two are the landscapes and office scenes on the Scyllarus Series A datasets [1]. The office scenes dataset contains 13 images whereas the landscapes imagery comprises 15 scenes. All three datasets have been acquired using hyper-spectral cameras with calibration charts so as to estimate the scene reflectance. For our experiments, we have computed the radiance maps using the sunlight spectrum acquired using a StellarNet spectrometer.

5.4.1 Results for Multiple Image Radiometric Calibration

We begin by illustrating the utility of our method to recover HDR imagery from multiple LDR images. For purposes of comparison, we have used the methods of Debevec and Malik [1997], Mitsunaga and Nayar [1999], and Robertson et al. [2003]. These methods have been used in many applications for radiometric calibration and are widely cited in literature. It is worth noting that the parameters for the method of Mitsunaga and Nayar [1999] have been set via cross validation for each of the three datasets. This contrasts with our method, which is devoid of free parameters.

To test the methods under study on different camera settings, LDR images were generated for each dataset using different spectral sensitivities and CCD response curves. For the Foster dataset, LDR images were rendered using the spectral sen-

[1] Accessible at http://scyllarus.research.nicta.com.au/data/

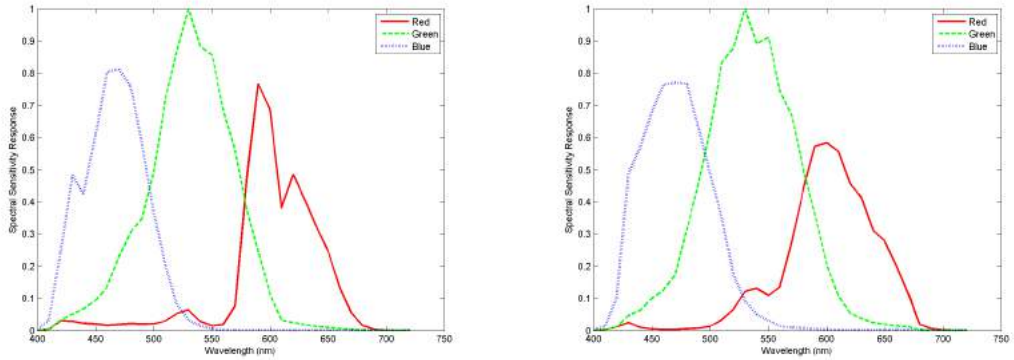

Figure 5.1: Spectral sensitivity of the Nikkon D80 (left) used on the Foster dataset; Cannon 20D spectral sensitivity (right) used for the Scyllarus datasets.

sitivity function of the Nikon D80 [2] and the camera response function of an Optura 981-114 CCD[3]. For the Scyllarus datasets, LDR images were rendered with a Canon2D spectral sensitivity function and a Optura 981-111 CCD camera response function. These spectral sensitivity functions and the camera response functions are shown in Figure 5.1 and Figure 5.2 respectively .

Use of these imagery, camera responses and color matching functions is important, since it permits the computation of both, LDR images for our experiments and radiance maps for quantitative evaluation. Figure 5.3 shows the ground truth radiance of a sample scene from the Foster dataset along with images generated at different exposure times. In the figure, the left-hand column shows the ground truth radiance maps, whereas images from the second column onwards correspond to the five LDR images, obtained by setting exposure time of 1/250, 1/125, 1/60, 1/30, and 1/15 seconds, respectively. Other parameters, including gain and aperture are assumed to be constant.

We generated radiance maps and camera response curves using our method and the alternatives with different number of input LDR images for all the scenes of both data sets. Figure 5.4 shows the results yielded by our method and the alternatives for the scene shown in Figure 5.3 when two (1/250 and 1/15) images are used. In the first column of Figure 5.4, we present the radiance images generated by the

[2] Accessible at http://www.cis.rit.edu/jwgu/research/camspec /db.php
[3] Accessible at http://www.cs.columbia.edu/CAVE/software /softlib/dorf.php

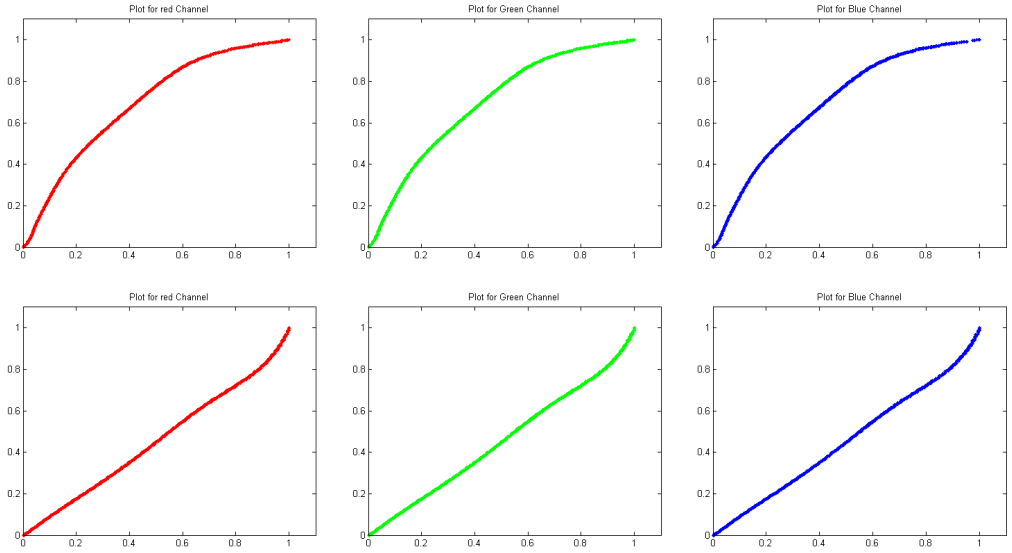

Figure 5.2: Camera response functions used to render our LDR imagery. Top row: Optura 981-114 CCD; Bottom row: Sony DX-9000- CCD. From left-to-right: Red, green, and blue channels.

Figure 5.3: Radiance and LDR images for a sample scene used in our experiments. From left-to-right: Ground truth normalised radiance image and LDR images generated using exposure time of 1/250, 1/125, 1/60 , 1/30 and 1/15 seconds, respectively.

methods. The second column presents tone mapped version of the radiance images. We use the same tone mapping operator with equal parameters for all methods. In order to qualify results, we take the normalized absolute difference of ground truth tonemapped images and those generated by each method. We show these error images for each method in the third column of Figure 5.4. These error images are represented on a red to blue map where red represents 1 and blue represents 0.

It can be observed from the error images, that for two input images our method result yields the least error, whereas Debevec's method yields the highest when two images are used. This is also observed from the camera response functions generated in Figure 5.5. Each row in Figure 5.5 represents result of a method, whereas the columns from left to right represent the resultant response function for red, green and blue channels. The black curve in each plot accounts for the ground truth camera response function used to render the imagery. It can be observed that for two input images our method approximates the camera's response function better than other methods do. Thus the performance of error images in Figure 5.4 can easily be correlated with the approximation of the camera response function in Figure 5.5.

Similar to Figure 5.4 in Figure 5.6, we show the resultant images and error images when three input images (1/250s, 1/60s and 1/15s) are used. It can be observed from the error images, that the overall error has been reduced by increasing the input LDR images. It can also be observed from the error images that our method outperforms others. This can also be validated from the camera response functions estimated by each method for three input images shown in Figure 5.7. The CRF plots in the figure show our method approximates ground truth CRF better than other methods when three input images are used.

In Figure 5.8, we show radiance, tone mapping and error images of the methods generated using 5 input images (1/250s, 1/125s,1/60s,1/30s and 1/15). It can be observed that the radiance, tone mapped and error images of all methods are highly similar. The similarity is due to reduced error of the of radiance images from ground truth radiance. This also suggest that around five images all the methods have approximated the ground truth radiance. This phenomenon can also be observed in the CRF plots shown in Figure 5.9, where the CRF plots of all methods are in good

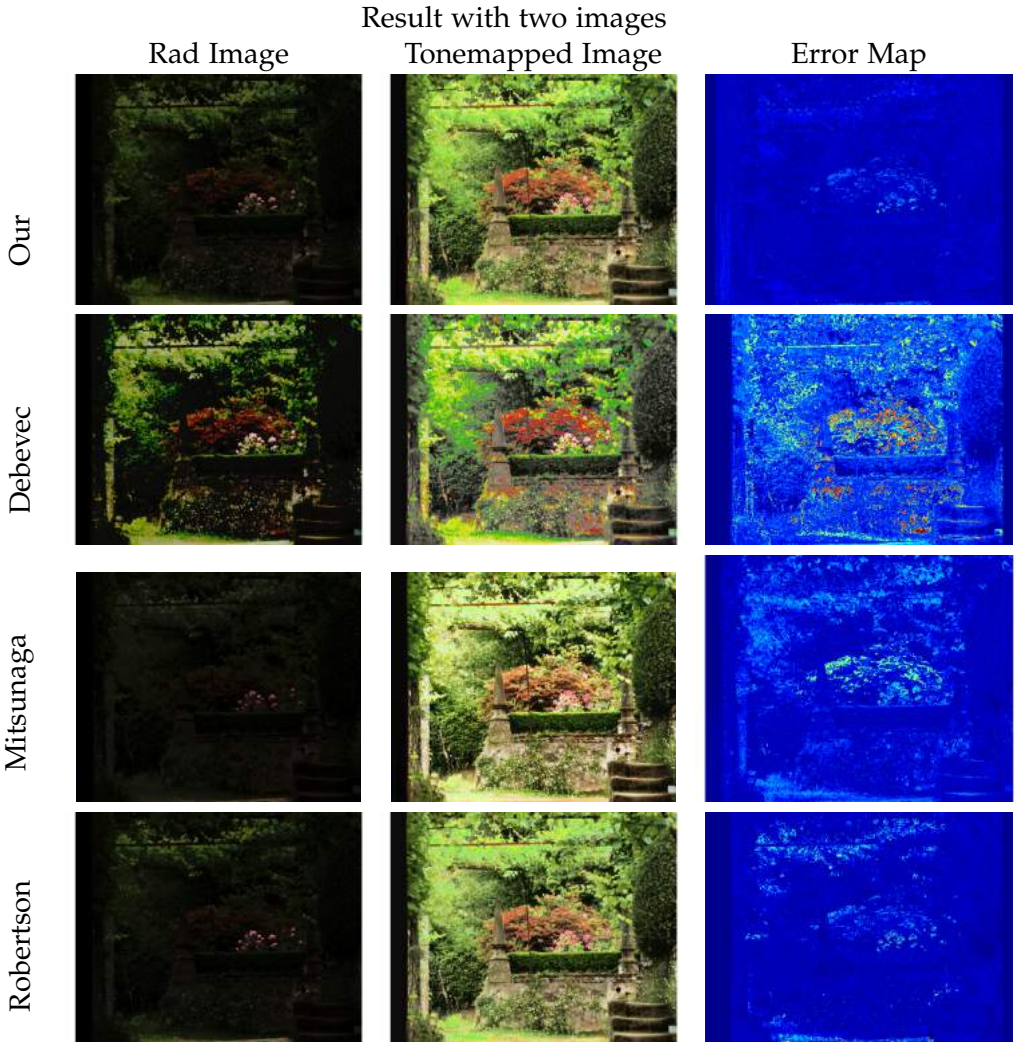

Figure 5.4: Radiance images , tonemapped radiance image and the error image generated by methods using two images for the sample scene on the Foster dataset shown in Figure 5.3. From top-to-bottom: Results delivered by our's, Debevec's , Mitsunaga's and Robertson's methods. The error images are generated using the ground truth and methods generated image. Error images are normalized where red represents 1 and blue represents 0

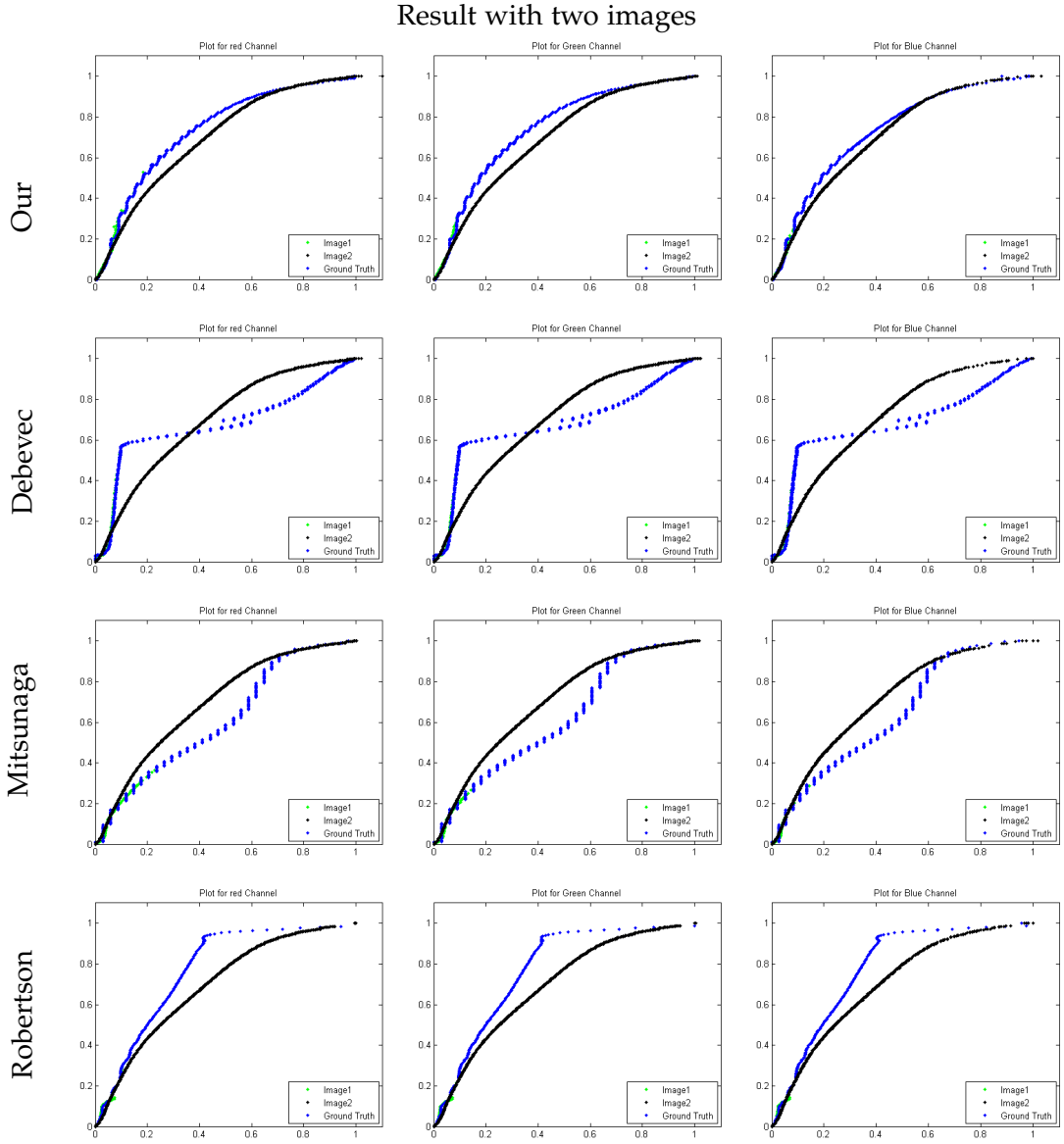

Figure 5.5: Camera response functions calculated using two images for a sample scene on the Foster dataset shown in Figure 5.3. From top-to-bottom: Resulsts delivered by our's, Debevec's , Mitsunaga's and Robertson's methods. The plots from left to right represent the red, green , and blue channels. The black curve in all the plots accounts for the ground truth camera response used to render the imagery.

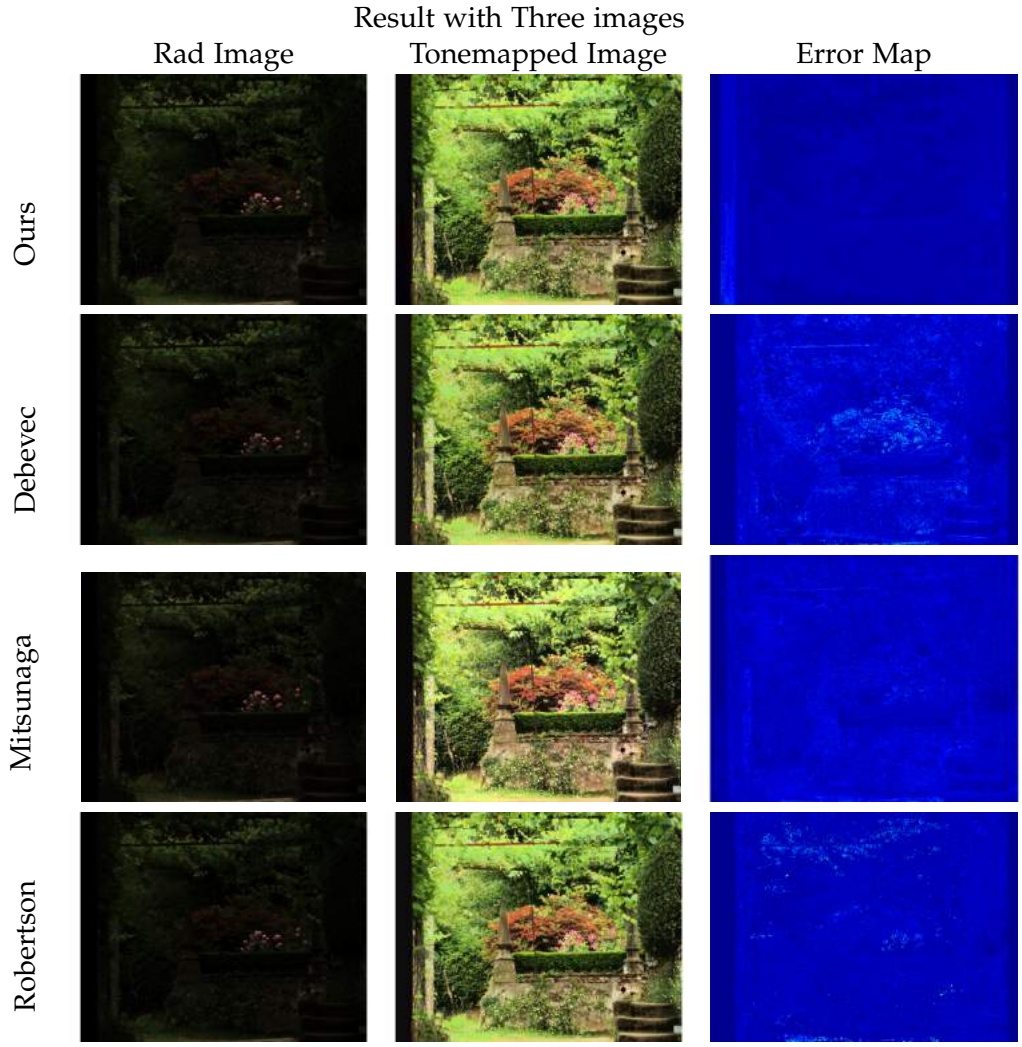

Figure 5.6: Radiance images , tonemapped radiance image and the error image generated by methods using three images for the sample scene on the Foster dataset shown in Figure 5.3. From top-to-bottom: Results delivered by our's, Debevec's , Mitsunaga's and Robertson's methods. The error images are generated using the ground truth and methods generated image. Error images are normalized where red represents 1 and blue represents 0

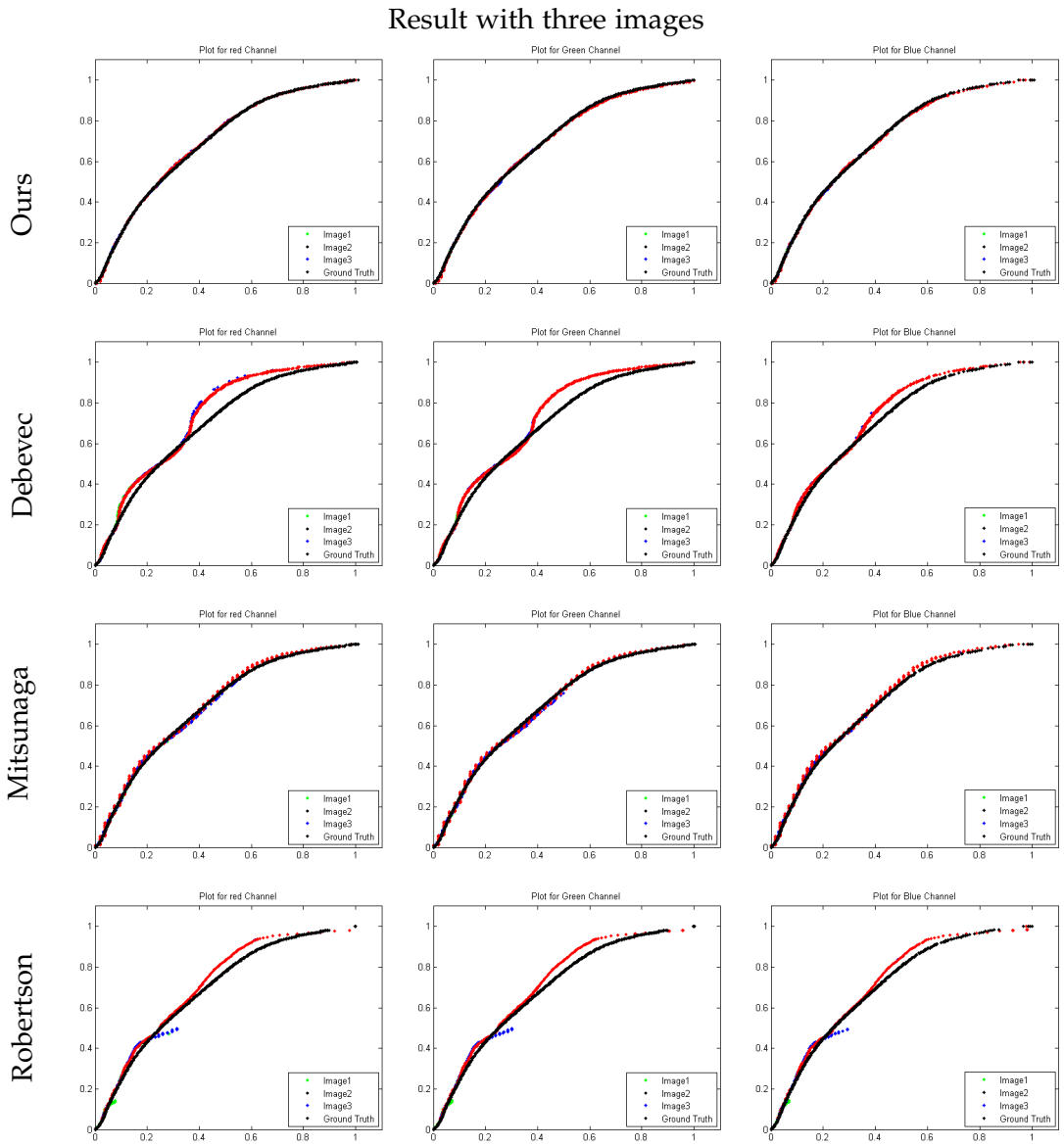

Figure 5.7: Camera response functions calculated using three images (1/250s, 1/60s and 1/15s) for a sample scene on the Foster dataset shown in Figure 5.3. From top-to-bottom: Resulsts delivered by our's, Debevec's , Mitsunaga's and Robertson's methods. The plots from left to right represent the red, green , and blue channels. The black curve in all the plots accounts for the ground truth camera response used to render the imagery.

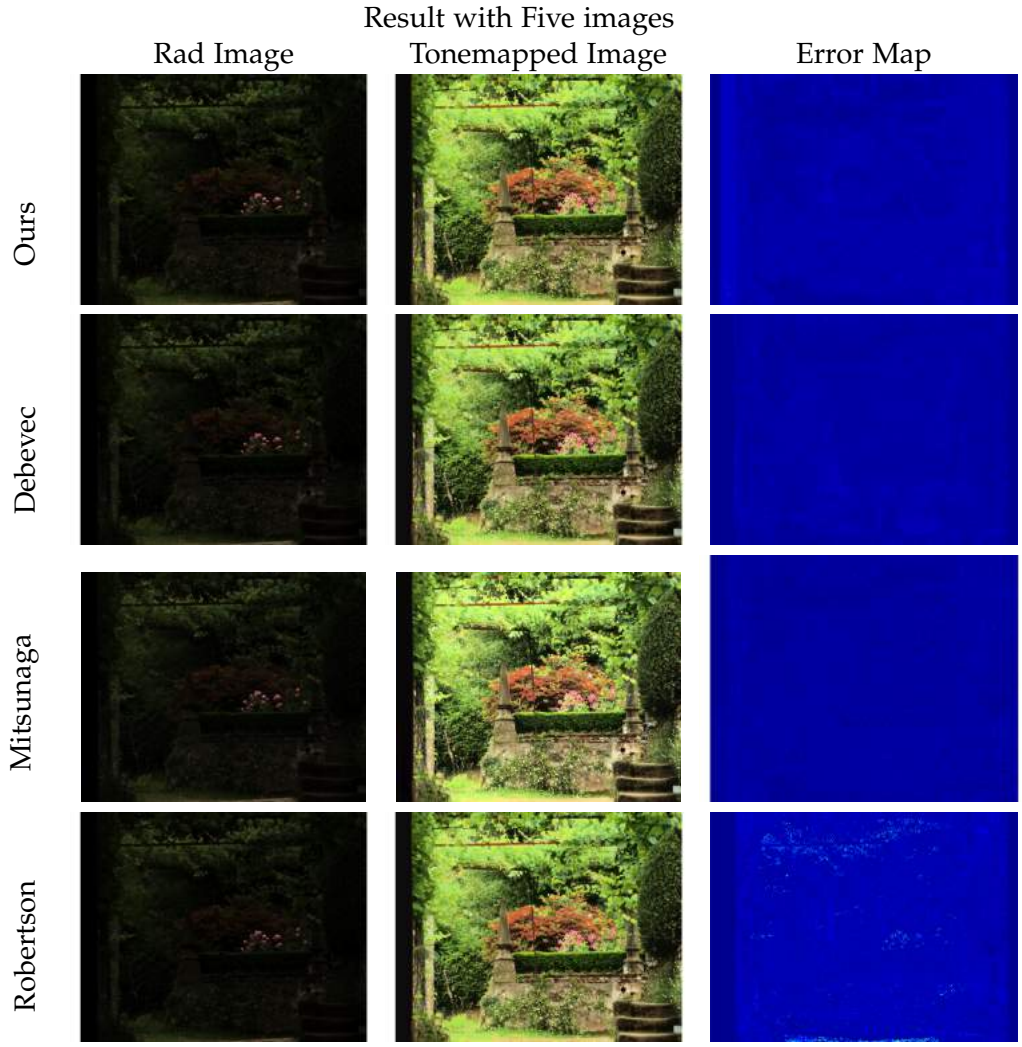

Figure 5.8: Radiance images , tonemapped radiance image and the error image generated by methods using three images for the sample scene on the Foster dataset shown in Figure 5.3. From top-to-bottom: Results delivered by our's, Debevec's , Mitsunaga's and Robertson's methods. The error images are generated using the ground truth and methods generated image. Error images are normalized where red represents 1 and blue represents 0

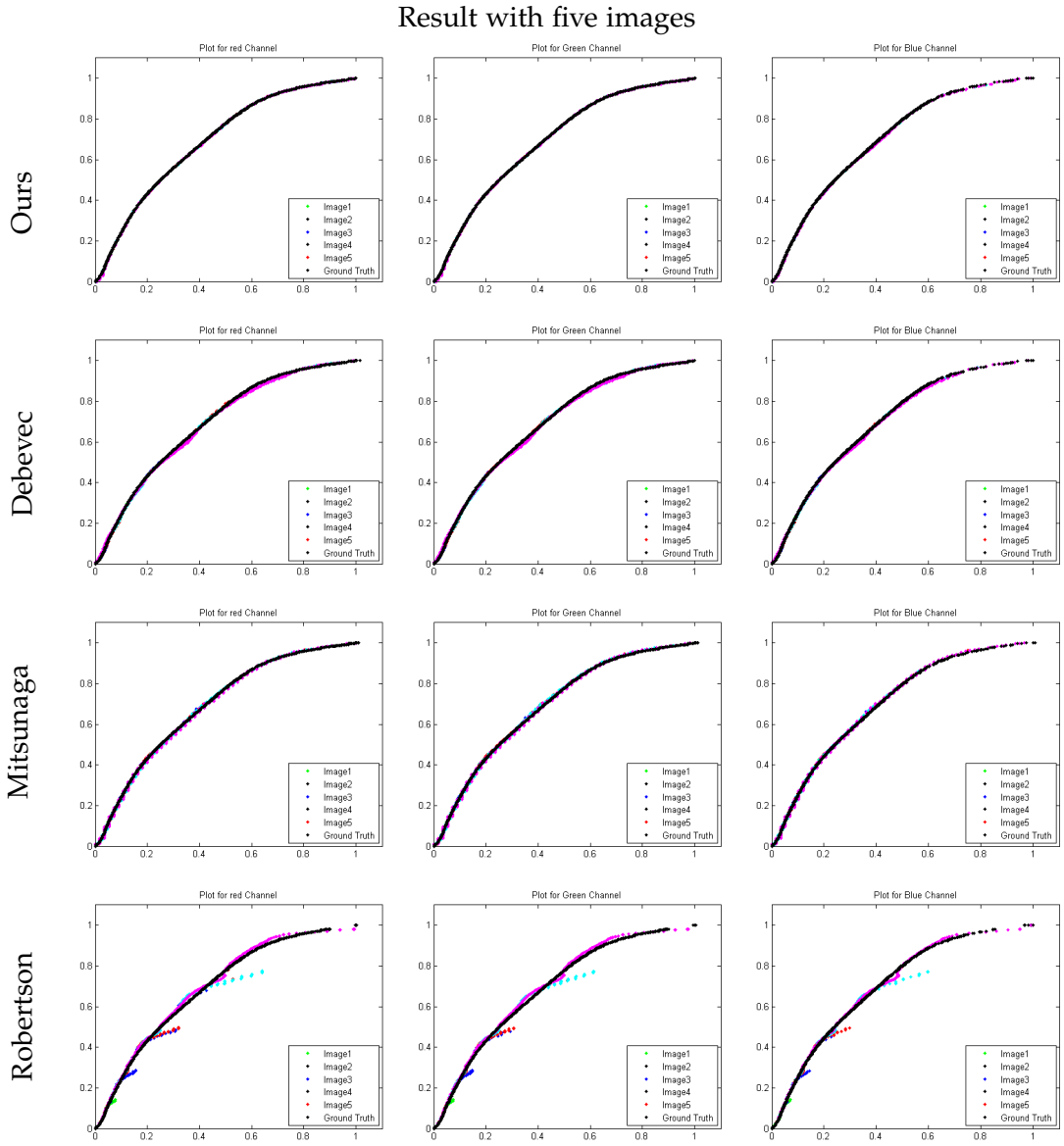

Figure 5.9: Camera response functions calculated using five images for a sample scene on the Foster dataset shown in Figure 5.3. From top-to-bottom: Results delivered by our's, Debevec's , Mitsunaga's and Robertson's methods. The plots from left to right represent the red, green , and blue channels. The black curve in all the plots accounts for the ground truth camera response used to render the imagery.

		Number of Images			
		1	2	3	5
Foster	Ours	**0.0693 ± 0.0175**	**0.0526 ± 0.0293**	**0.0086 ± 0.0069**	**0.0022 ± 0.0011**
	Debevec	NA	0.1110 ± 0.0240	0.0612 ± 0.0474	0.0075 ± 0.0059
	Mitsunaga	NA	0.0695 ± 0.0333	0.0092 ± 0.0043	0.0037 ± 0.0018
	Robertson	NA	0.0642 ± 0.0382	0.0410 ± 0.0367	0.0252 ± 0.0284
Landscapes	Ours	**0.0593 ± 0.0253**	**0.0334 ± 0.0120**	**0.0153 ± 0.0128**	**0.0100 ± 0.0078**
	Debevec	NA	0.2318± 0.1089	0.0441 ±0.0271	0.0165 ± 0.0076
	Mitsunaga	NA	0.1038 ± 0.0263	0.0205 ± 0.0107	0.0152 ± 0.0208
	Robertson	NA	0.1236± 0.0253	0.0696 ± 0.0260	0.0670 ± 0.0415
Offices	Ours	**0.1096 ± 0.0278**	**0.0313 ± 0.0104**	**0.0138 ± 0.0077**	**0.0115 ± 0.0059**
	Debevec	NA	0.1560± 0.0251	0.0582 ± 0.0356	0.0178 ± 0.0059
	Mitsunaga	NA	0.1247 ± 0.0287	0.0383 ± 0.0236	0.0164 ± 0.0074
	Robertson	NA	0.1178 ± 0.0251	0.0711 ± 0.0234	0.0694 ± 0.0323

Table 5.1: RMS error and standard deviation across each of the datasets under study for ours, Debevec's, Mitsunaga's, and Robertson's methods.

accordance the ground truth.

In Figure 5.10, we show the comparison of CRF plots of the methods for two and three images. Since for five input LDR images the CRF plots of all the methods were significantly close to the ground truth, we stick to two and three images for comparison. The left-hand panel of Figure 5.10 shows the result for two input LDR images, whereas the right-hand panel shows the result for three input LDR images. In both panels we show the red, green and blue channels from top to bottom, respectively. It can be observed that our method outperforms for any number of input LDR images.

Now, we turn our attention to a more quantitative analysis. In Table 5.1, we show the mean RMS and its standard deviation for each of the alternatives for the cases where one, two, three, and five LDR images are used at input. Using the given methods we calculate the average of RMS of images for each dataset. In case of multiple images, our method outperforms the others for all three datasets. The standard deviations also suggest our method to be more stable in comparison with the alternatives.

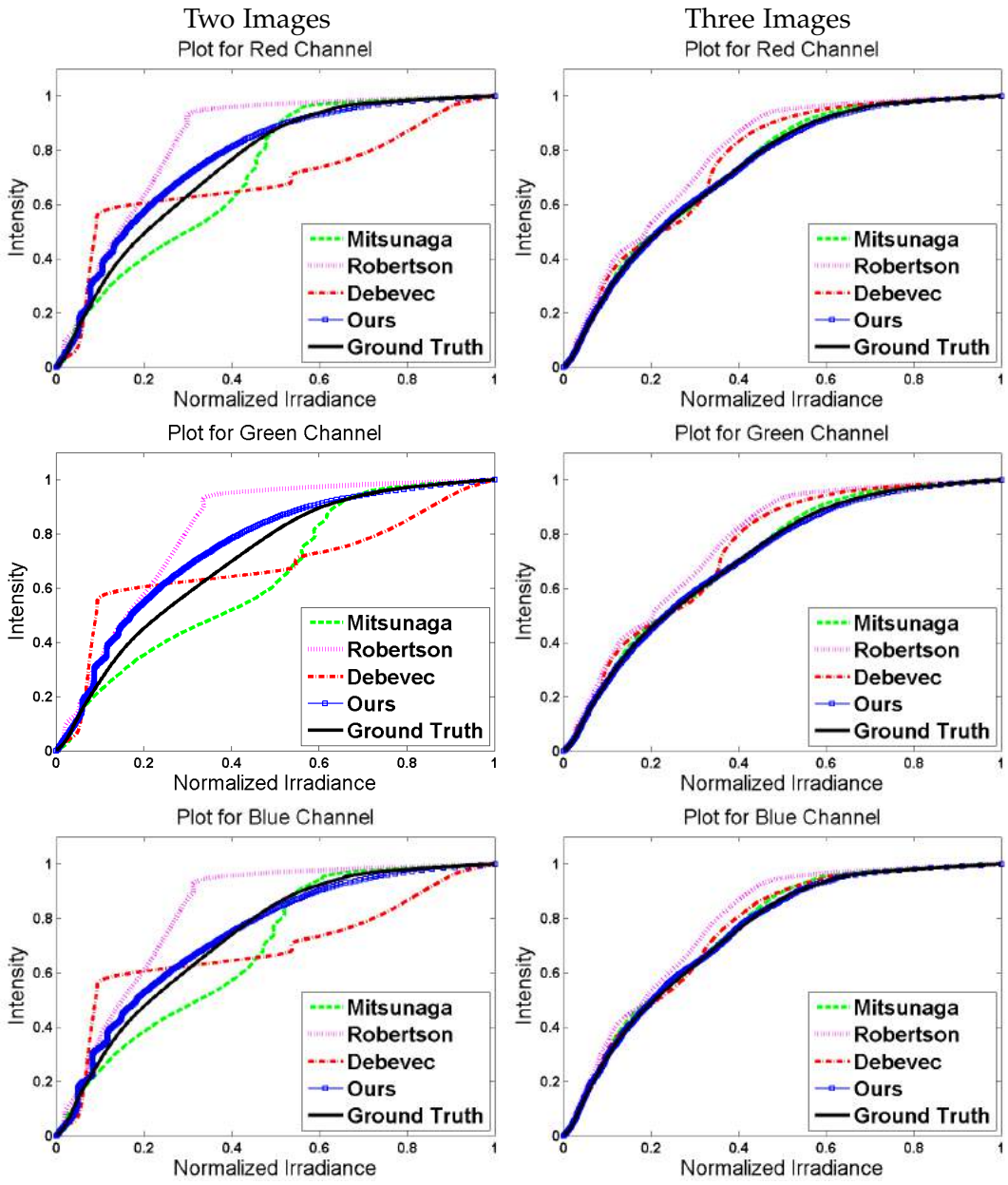

Figure 5.10: Left-hand column: Camera response function for the Optura 981114 CCD recovered by our method and the alternatives when two LDR images are used; Right-hand column: Camera response function recovered when three LDR images are used at input. The rows, from top-to-bottom show the functions for each of the colour channels, *i.e.* red, green and blue.

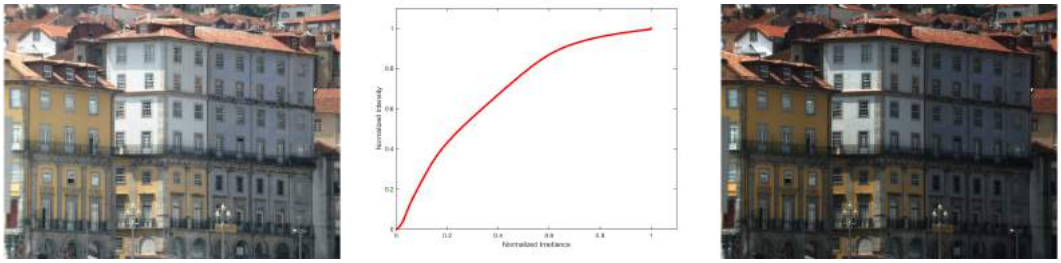

Figure 5.11: Left to right: Input LDR image for SIRC, Groundtruth camera response curve, Groundtruth irradiance image taken from CRF and input LDR image

5.4.2 Results for Single Image Radiometric Calibration

For SIRC, we show the input LDR image in Figure 5.11. This input image was taken at the exposure time of 1/60 seconds which contains a balance of bright and dark regions. This input LDR image was captured with the camera response function shown in the middle of Figure 5.11. The rightmost image in Figure 5.11 shows the groundtruth irradiance image calculated using the inverse camera response function applied to the input LDR image. In Figure 5.12, we show at each row, the results of our method at different iterations. Along the column, we show for each iteration its irradiance image, intermediate irradiance image, and the CRF plots of red, green and blue channels. In the CRF plots, ground truth and method generated CRF are shown in black, and green respectively. In the first row results for first iteration are shown. In the second row, the results are shown when the method reaches middle point of each iteration. In the third row, the output at the final iteration is shown. With each iteration the irradiance image, intermediate image, and CRF plots gets updated to approximate the ground truth as shown in the Figure 5.11.

For a quantitative analysis, we have used the RMS error for the recovered irradiance images as compared to the ground truth irradiance images for each image in the three datasets. In Table 5.1, we show the mean RMS and its standard deviation for our method when one LDR image is used at input. Since only our method can work with single image, we have used not applicable (NA) abbreviation for other methods for single image. Note from Table 5.1, that the error with single image is higher than multiple case, and it decreases as the number of input images increases.

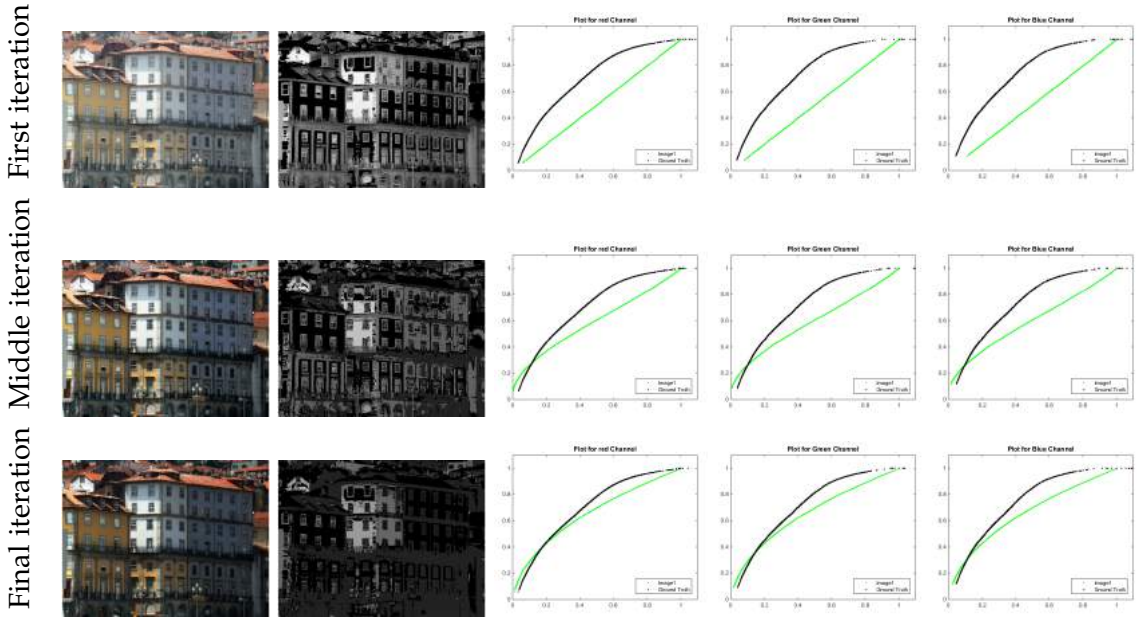

Figure 5.12: Irradiance image and camera response function for the Optura 981114 CCD recovered by our method when using single LDR image. Top to bottom: First row shows the irradiance image and CRF of red, green and blue channel at first iteration. Second row shows the irradiance image and CRF at the middle of iteration. Bottom row shows the result at final iteration.

5.4.3 Results on Panorama Images

In this sub section we illustrate the utility of our method for panorama images. To compare with other methods we use two input images for each scene to recover radiance images. These radiance images were tonemapped using Larson et al. [1997] and then stitched to form the panorama. In Figure 5.13 we show sample results on panorama construction using HDR images obtained by our method and the alternatives. In the figure, we show a panorama comprised of four tone mapped HDR images stitched together using the Microsoft Research Image Composite Editor (ICE) [4]. In Figure 5.14 we mark the regions of interest and show them in detail by cropping and magnifying them, in the second and third row of the figure. By observing the sunlit and shadowed regions on the pavement and the wall in the magnified images, it becomes evident that our method preserves better the finer details and texture than the other approaches. This also implies that our method has better estimate of

[4] research.microsoft.com/en-us/um/redmond/groups/ivm/ice/

radiance image, which when tone mapped, displays finer details in the scene.

In Figure 5.15, we show another panorama of landscape stitched using four scenes. Our approach delivers a more realistic landscape. Moreover, even when only two LDR shots per HDR image are used, the lake panorama delivered by our method still depicts a fair amount of detail on the water and better preserves the contrast on the background tree line. These would appear blurry with the inclusion of further shots. In Figure 5.15 the result of Robertson et al. [2003] shows over-saturation around the sun. This effect is due to the CRF estimate of Robertson method which reaches maximum intensity at the mid of the normalized irradiance.

5.5 Time Complexity of the algorithm

Assuming we have total M images, C channels, U pixels in one channel, B brightness levels in image, and an average of K number of pixels sharing the same brightness levels, the average case running complexity of one iteration is $\mathcal{O}(UCM + BCK)$. Also for foster data set average number of iterations were 4.2, and for Sycllarus data sets it took on average 5.3 iterations.

5.6 Summary

In this chapter, we have presented a general radiometric calibration method for HDR imaging which is devoid of free parameters and can deliver reliable imagery with as few as two LDR images as well as with single well exposed image. The method presented here is based upon an optimisation approach which can be solved efficiently using the Lq Weiszfeld algorithm in Aftab et al. [2015]. We have shown how the initial estimate of the radiance map (or irradiance image for single image case) can be recovered for initialisation purposes. We have also presented quantitative and qualitative results making use of ground truth radiance maps and irradiance images and illustrated the utility of our method for the creation of real-world panoramas. Throughout our experiments we have compared our results with those yielded by alternatives elsewhere in the literature. In the next chapter, we shall present our

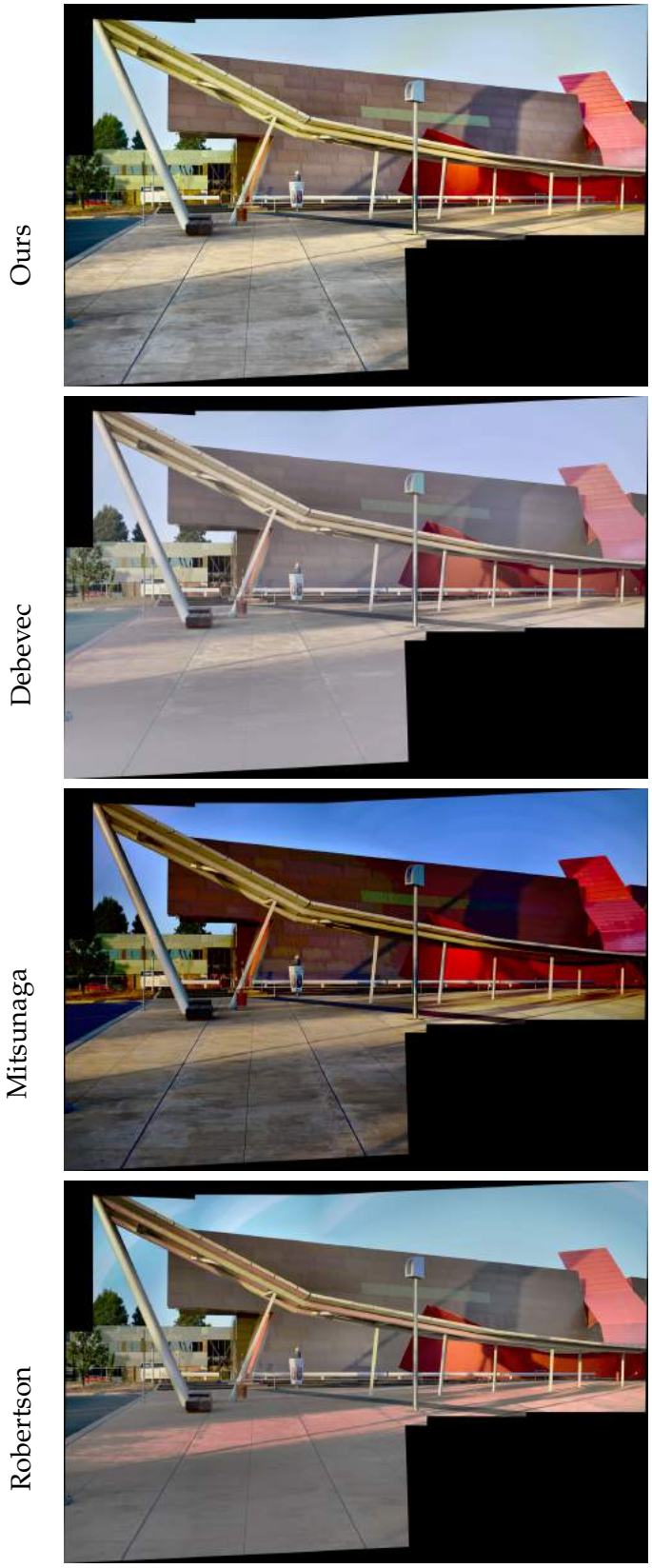

Figure 5.13: Panoramas generated by stitching HDR tone-mapped images delivered by ours, Debevec's , Mitsunaga and Robertson's method.

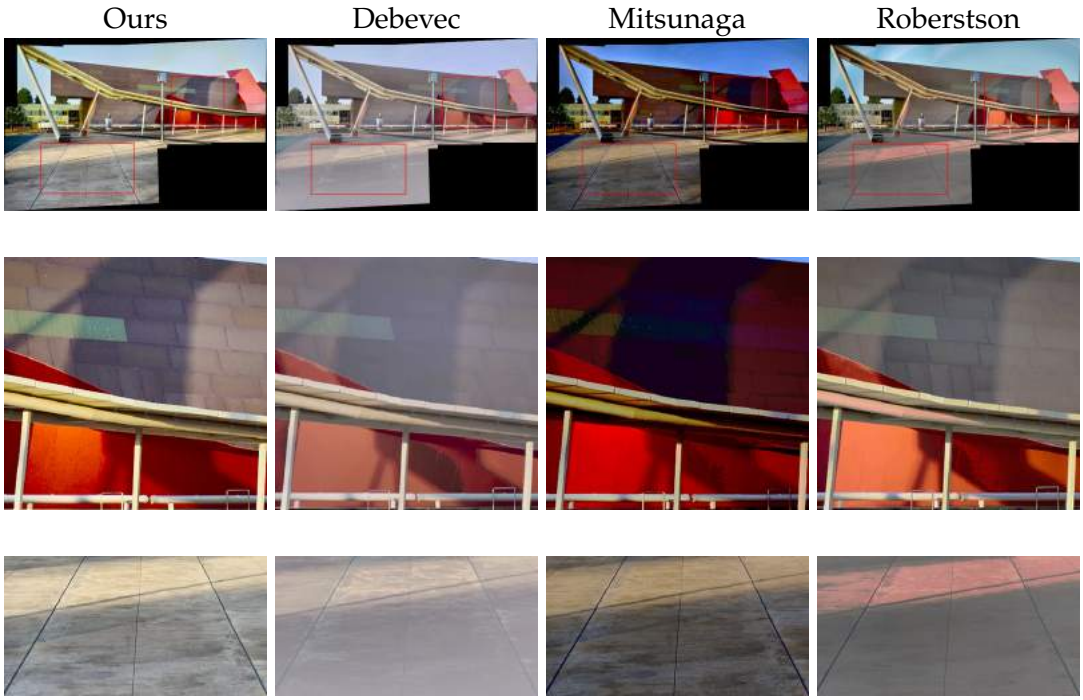

Figure 5.14: Panoramas generated by stitching tone-mapped HDR views using 2 LDR images. (From left to right) Results yielded by our's (1st column) Debevec's (2nd column); Mitsunaga's (3rd column) and Robertson's (4th column) methods.First row shows the generated results ; second and third row shows the cropped regions magnified for respective method

method to compress the dynamic range of the generated radiance/high dynamic range images, in order to display them on low dynamic range displays. Our presented compression method ensures to preserve the finer details of radiance images.

Figure 5.15: Panoramas generated by stitching HDR tone-mapped images delivered by ours, Debevec's , Mitsunaga and Robertson's method.

Chapter 6

High dynamic range compression

In this chapter, we present a method for dynamic range compression of a radiance image. Our proposed method combines the effects of global and local techniques for dynamic range compression. Initially, it uses a global technique to generate a number of low dynamic range (LDR) images from the input high dynamic range (HDR) image. It then uses a localized technique based on the contrast at a neighborhood of each pixel for the input LDR images, to minimize an energy function for graph regularization. This approach generates a dynamic range compressed image that accounts for the local contrast sensitivity of human visual system (HVS) as well as providing a consistent global approach to generate a series of initial LDR images. In this chapter, we also illustrate the utility of our method by generating the compressed dynamic range image and compare qualitatively and quantitatively with the results of methods widely used in literature and used professionally. Our quantitative comparison is based on three aspects of dynamic range compression, these are 1) loss of contrast 2) amplification of contrast, and 3) reversal of contrast.

In Section 6.1, we explain our approach for compressing the dynamic range of a radiance image by using graph regularization on multiple generated low dynamic range images. In Section 6.2, we explain our approach of exposure selection for generating multiple LDR images. In Section 6.3, we show the qualitative and quantitative comparison of results of our method with those widely used in literature as well as those used professionally. Finally, in Section 6.4 we summarize our work presented in this chapter.

6.1 Method

We consider the problem of dynamic range compression as that of fusing of images taken at different exposures. Since the resolution of gradient for brightness is observed differently at different image exposures, an HDR image can hence be used to generate a set of low dynamic range images which are mapped using a camera response curve. The general equation for this mapping is given as:

$$I_i(u) = f[I_{HDR}(u)T_i] \tag{6.1}$$

here, f is the camera response function which is a camera specific mapping function of image intensity $I(u)$ at the pixel location u, to the radiance map $I_{HDR}(u)$ scaled by exposure time T of the image i for channel c. Using Equation 6.1 with radiance map $I_{HDR}(u)$ in hand, we apply a reverse model to generate a number of low dynamic range images given as:

$$\widehat{I}_i(u) = \beta I_{HDR}(u)T_i \tag{6.2}$$

β is a constant introduced for assumption of a linear response function.

Here, we generate LDR images by changing the exposure time, and bracketing the resultant maximum value to the maximum of the compressed dynamic range. We also assume a linear response function for this operation. Assuming multiple exposed LDR images are in hand, the problem of compressing a high dynamic range image to a low dynamic range one can be considered as of obtaining a continuous labeling of linear image pixels obtained by exposure selection technique. These image labels are then used to combine these bracketed LDR images to obtain a final compressed dynamic range image. Here, we use a method akin to image labeling techniques which employ energy minimisation for Markov Random Fields with a known prior in multi-region segmentation tasks (Boykov and Jolly [2001]).

Consider a set of n LDR images $\mathcal{I} = \{I_1, I_2, \ldots, I_n\}$ generated by bracketing the I_{HDR} image at n different exposure times. We can then define the pixel u in the

compressed dynamic range image $\widetilde{I}$ as follows

$$\widetilde{I}(u) = \sum_{i=1}^{n} x_{u,i} \widehat{I}_i(u) \tag{6.3}$$

where $\widehat{I}_i(u)$ denotes the pixel u in the ith image in $\mathcal{I}$, $x_{u,i}$ is a continuous label in the interval $[0, 1]$ such that

$$\sum_{i=1}^{n} x_{u,i} = 1 \tag{6.4}$$

To recover the labels $x_{u,i}$, we adopt a graph regularisation approach which is formulated as the following energy minimisation problem

$$\min c(X) = \lambda \sum_{\substack{I_i \in \mathcal{I} \\ u \in \widetilde{I}}} d^2(x_{u,i}, K \otimes B_{u,i}) + \sum_{\substack{I_i \in \mathcal{I} \\ u \in \widetilde{I} \\ u \sim v}} w_{u,v} d^2(x_{u,i}, x_{v,i}) \tag{6.5}$$

where d is a distance metric, $w_{u,v}$ is the affinity between pixels u and v, λ is a constant which weights the contribution of the first term in the right-hand side ; B is the lightness component of the L*a*b* color space of the image, and K is the kernel convolved with lightness component of the image to get the change in contrast of lightness in the image $\widehat{I}$. In Equation 6.5, $\otimes$ represents convolution of kernel K with B at location u. In the equation above, $\sim$ denotes a neighbourhood relation between pixels. Here, we employ a four-neighbourhood, where each pixel in the image is connected to two neighbouring pixels in the horizontal direction and two in the vertical direction.

Before continuing with our analysis, we can provide an intuitive interpretation of the cost function above. The first term in the left-hand side of Equation 6.5 aims to give higher label values to the pixels which have higher contrast change with respect to their neighbours. The kernel convolved with the lightness of the image is given as:

$$K = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} . \tag{6.6}$$

The kernel given above is used to give higher label value to the pixel whose lightness contrast is higher compared to the neighbours, this is because those areas of the image need more contribution where resolution of contrast gradient is higher to avoid missing the image details. Thus the pixels expected to have higher contrast will contribute more to $\widetilde{I}(u)$. Similarly, kernels for edge detection can also be used for this purpose. The second term enforces spatial smoothness across the image lattice. As a result, for the affinity function $w_{u,v}$, we use the expression

$$w_{u,v} = \begin{cases} exp(-||u-v||^2/2\sigma^2) & v \sim u \\ 0 & otherwise \end{cases} \tag{6.7}$$

where σ is set constant.

Furthermore, note that Equation 6.5 can be rewritten as follows

$$\min c(X) = \lambda \sum_{u \in \widetilde{I}} d^2(\mathbf{x}_u, \mathbf{b}_u) + \sum_{\substack{u \in \widetilde{I} \\ u \sim v}} w_{u,v} d^2(\mathbf{x}_u, \mathbf{x}_v) \tag{6.8}$$

where $\mathbf{x}_u = [x_{u,1}, x_{u,2}, \ldots, x_{u,n}]^T$ is a n-dimensional label vector and $\mathbf{b}_u = [K \otimes B_{u,1}, K \otimes B_{u,2}, \ldots, K \otimes B_{u,n}]^T$ is an n-dimensional vector representing convolution of kernel K with brightness intensity $B_{u,i}$ of i^{th} bracketed image.

It is worth noting that the minimisation problem above is akin to other energy minimisation functionals comprised of a data term, which enforces the consistency between the calculated and given labels, and a smoothness term, which penalizes disparate labellings for neighbouring pixels. With the L2-norm, the cost function in Equation 6.8 becomes

$$\min c(X) = \lambda \sum_{u \in I_{HDR}} ||\mathbf{x}_u - \mathbf{b}_u||_2 + \sum_{\substack{u \in I_{HDR} \\ u \sim v}} w_{u,v} ||\mathbf{x}_u - \mathbf{x}_v||_2 \tag{6.9}$$

Since the label values are constraint to assume values in the range $[0, 1]$, this leads to a convex quadratic optimisation problem akin to the Random Walks segmentation framework proposed in (Grady [2006]), where the above cost function can be re-

written in the following matrix

$$\min f(\mathbf{X}) = tr(\mathbf{X}^T\mathbf{L}\mathbf{X}) + \lambda||\mathbf{X} - \xi\mathbf{1}||^2 \tag{6.10}$$

where $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]^T$ is a N by n matrix of label vectors for the N pixels in the compressed LDR image and $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is the combinatorial Laplacian (Chung [1997]) computed using the adjacency matrix $\mathbf{W}$ with entries $w_{u,v}$ and the degree matrix $\mathbf{D}$, *i.e.* a diagonal matrix whose diagonal elements are given by the row-sums of $\mathbf{W}$. In the equation above, $\mathbf{1}$ is an N by n matrix whose entries are all unity.

The optimisation problem in Equation 6.10 is a convex one whose global minimum can be found by solving the following systems of linear equations

$$(\mathbf{L} + \lambda\xi^2\mathbf{1})\mathbf{X}_{.,i} = \lambda\xi\mathbf{1}^T\mathbf{Y}_{.,i}(i = 1, \ldots, n) \tag{6.11}$$

where $\mathbf{X}_{.,i}$ denotes the i^{th} column of $\mathbf{X}$.

Further, nonetheless the above equations can be solved with classical Gaussian-Siedel or conjugate gradient methods, we note that, due to the sparse nature of $\mathbf{L}$ and the fact that the equations share the same left-hand side, a more efficient approach is the application of sparse Cholesky factorisation (Davis [2006]). Thus, we first factorise $\mathbf{L} + \lambda\xi^2\mathbf{1}$ into the product of an upper-triangular matrix $\mathbf{Q}_h$ and its transpose $\mathbf{Q}^T$, which is a lower-triangular matrix. Then two trivial linear equations, $\mathbf{Q}\mathbf{H} = \xi\mathbf{1}^T\mathbf{Y}_{.,j}$ and $\mathbf{Q}^T\mathbf{X}_{.,j} = \mathbf{H}$ are solved via forward and backward substitution to recover $\mathbf{X}_{.,i}$. This is even more efficient since $\mathbf{Q}$ can be "recycled" and the sparse Cholesky factorisation only needs to be applied once for recovering multiple columns of $\mathbf{X}$.

6.2 Exposure Selection

In this section, we explain the exposure selection approach of our method. In order to have multiple exposed images, we need to find the number of images and their dynamic range such that each image represent a different set of contrast information from the HDR image. To achieve exposure selection various methods have been

proposed in the literature (Pourreza-Shahri and Kehtarnavaz [2015a]; Huang and Chiang [2013]; Hirakawa and Wolfe [2010]; Barakat et al. [2008]). For experiments and method comparison we adopted a simple linear approach for exposure selection. We need varying exposure values such that, for each exposure a different set of information is depicted. For this, first we make a histogram of the normalized gray-scale HDR image. The histogram is represented as probability distribution function containing M histogram bins. M is chosen large enough to cover the dynamic range of the image. The probability distribution function of the normalized gray scale image is represented as $PDF_m(I_{NGHDR})$. Here m refers to the bin index of the PDF and I_{NGHDR} represents the normalized gray scale HDR image. Using the probability distribution function we obtain the cumulative distribution function represented as $CDF_m(I_{NGHDR})$. We take C linearly distributed points between 0 and 1 as anchor points and find those points in the normalized CDF that have the minimum squared difference with respect to the anchor points. This squared difference is represented as

$$diff = (CDF_m(I_{NGHDR}) - A_l)^2 \tag{6.12}$$

here, m is the bin number given as $m = 1, 2, \ldots, M$ and A_l is the l^{th} anchor point where $l = 1, 2, \ldots, C$. From all m index of CDF, we find the ones that have the minimum $diff$ with A_l. It is possible that multiple indexes/bins have same $diff$ from the A_l. We chose only those anchor points that have a unique index/bin with minimum distance. Indexes having unique minimum distance suggest that the level of information they possess is not available with other indexes. Each of these unique indexes becomes a cap for HDR image with which an LDR image is generated. Thus resulting in a multiple LDR images exposed differently such that each represent a different set of image information.

6.3 Experiments

As discussed earlier, our method generates a number of LDR images from the input HDR image using the exposure selection method described above. In Figure 6.1, we

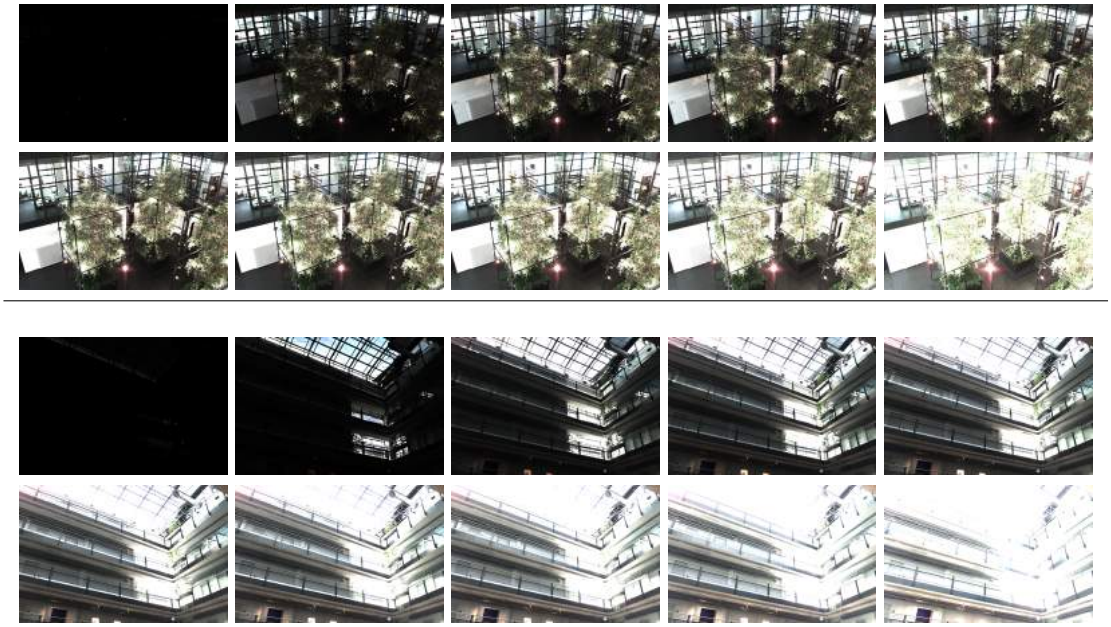

Figure 6.1: LDR images at different exposure obtained using exposure selection approach for two scenes: top panel first scene; bottom panel second scene

show the LDR images at different exposures generated by our exposure selection technique for two different scenes. Our exposure selection method generates 10 LDR images for each given scene. It can be observed that different level of detail is available at each image. These images when fused by our minimization approach results in a final combined LDR image which represents the overall details of the scene.

In Figures 6.2 and 6.3 we show the resultant images of our method at each step. In the left hand column of both left and right panels of Figures 6.2 and 6.3 we show LDR images of the two scenes shown in Figure 6.1. For each of these LDR images our method generates a contrast image by convolution of kernel given in Equation 6.6 with the lightness component of LDR L*a*b form. For each LDR image shown in Figures 6.2 and 6.3, we show their respective contrast images next to them. These contrast images are then used in the energy minimization with pixels neighborhood affinities to generate probability maps for each LDR image. These probability maps are used as image labels as represented in Equation 6.3. We show these probability maps for each given LDR images next to the contrast images shown in Figures 6.2

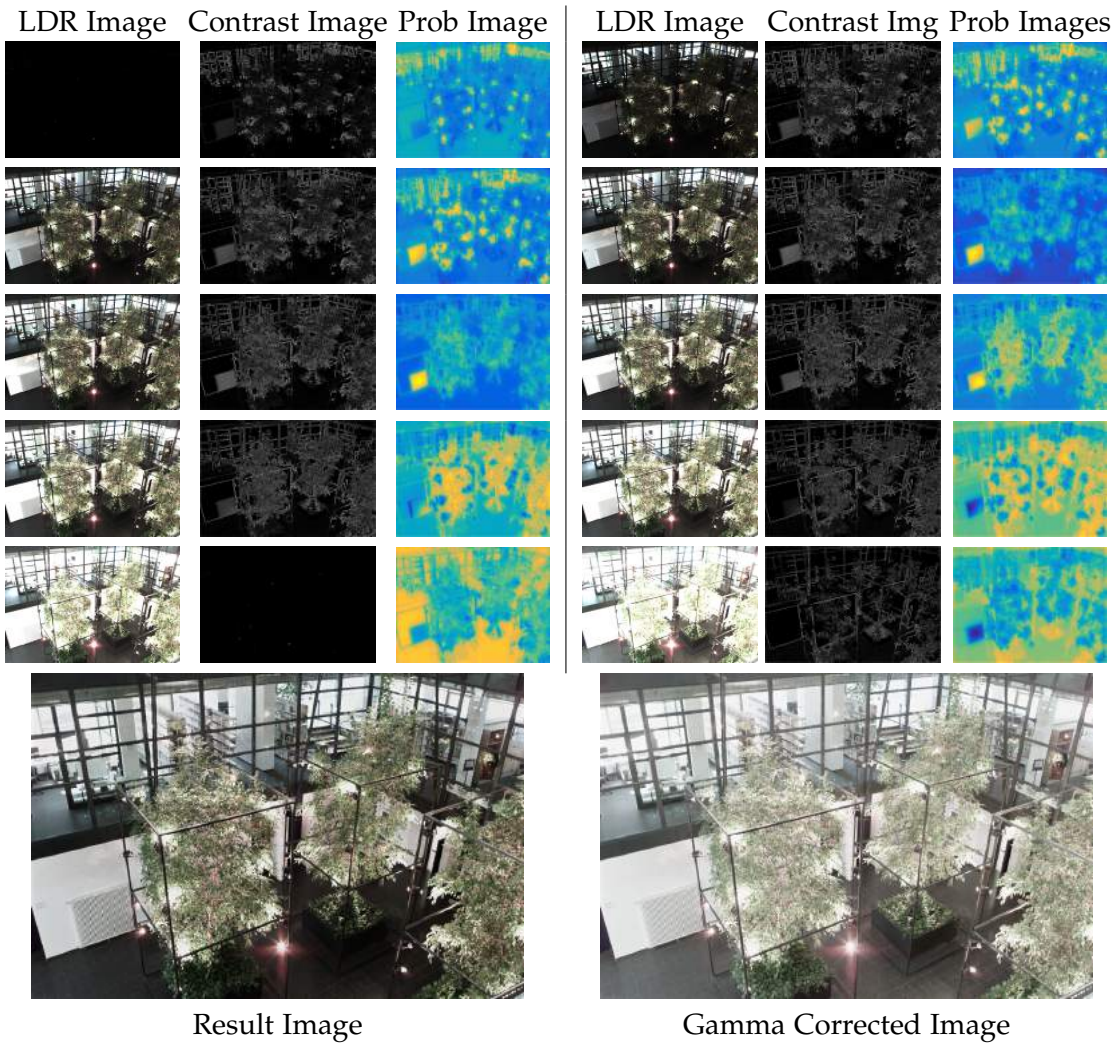

Figure 6.2: Contrast Images and probability images generated by our method using LDR images obtained by exposure selection. For both left and right panel the first column contains the LDR images generated by the exposure selection, second column contains the contrast images obtained from the LDR images and the third column represents the probability images generated from the corresponding contrast images. Last row shows the resultant compressed dynamic range image, and its gamma corrected image. Gamma used for correction is 2.2

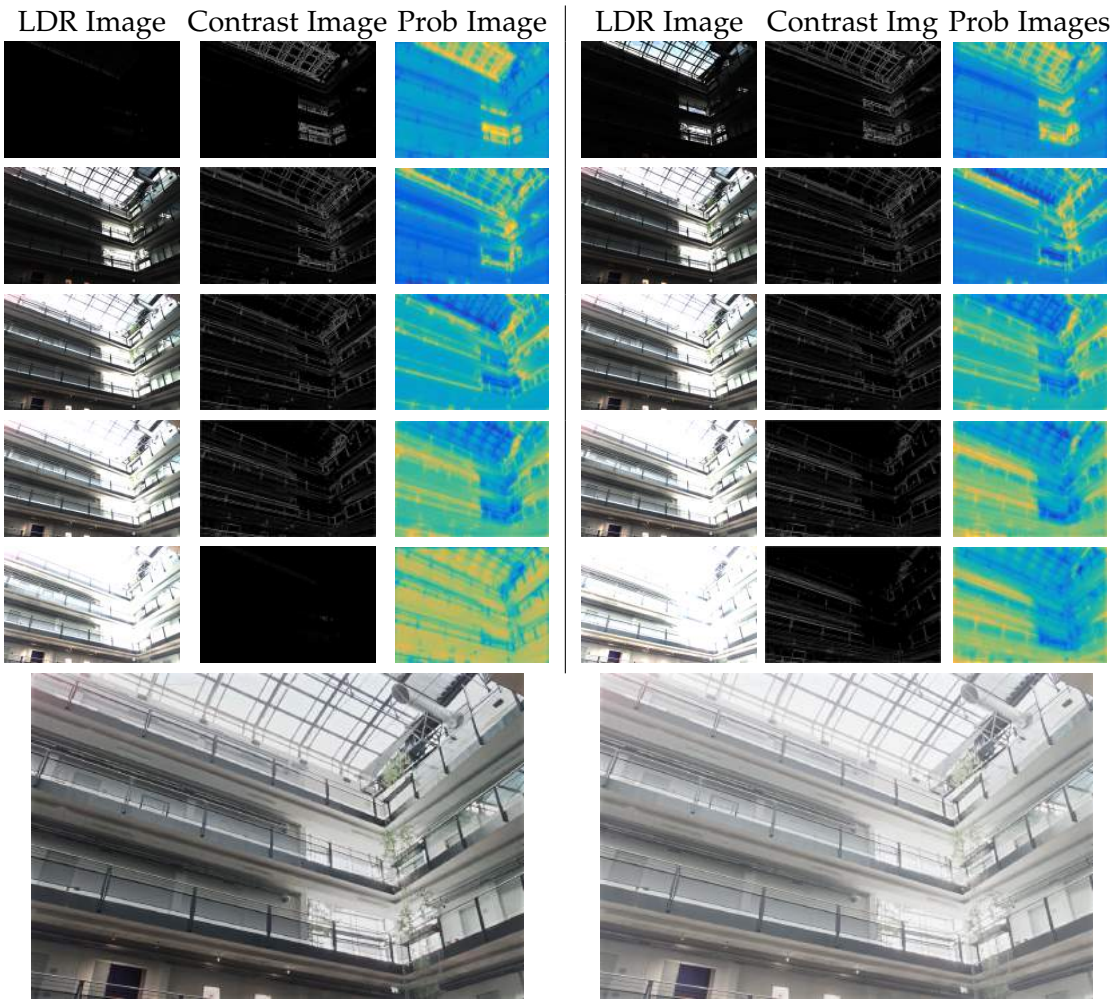

Figure 6.3: Contrast Images and probability images generated by our method using LDR images obtained by exposure selection. For both left and right panel the first column contains the LDR images generated by the exposure selection, second column contains the contrast images obtained from the LDR images and the third column represents the probability images generated from the corresponding contrast images. Last row shows the resultant compressed dynamic range image, and its gamma corrected image. Gamma used for correction is 2.2

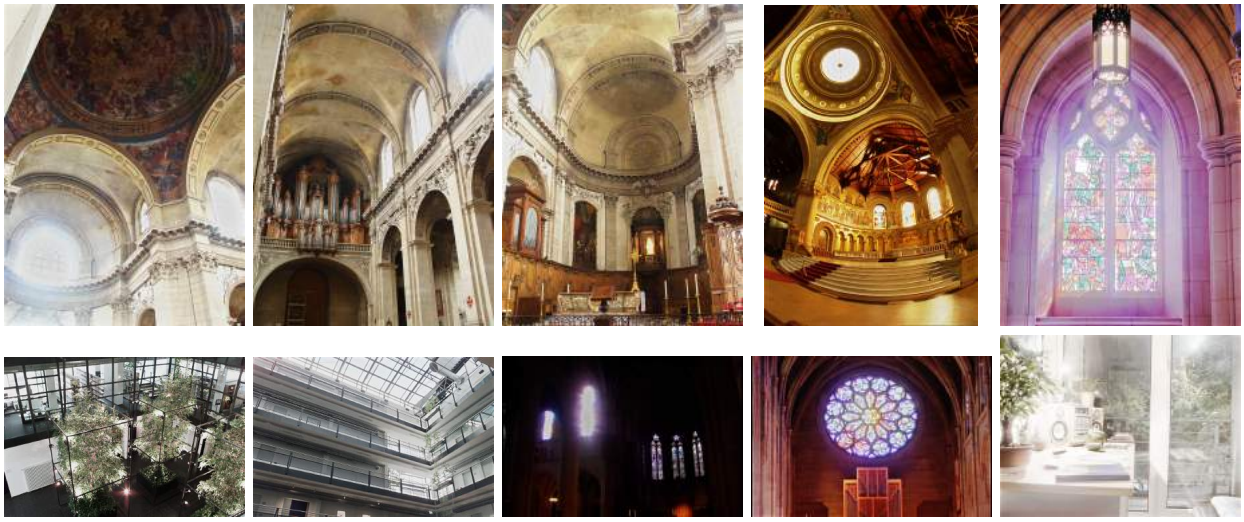

Figure 6.4: High dynamic range images widely used for method comparisons. We use these HDR images to test the dynamic range compression technique proposed in this paper in comparison with results of methods used professionally and cited in literature.

and 6.3. The resultant image generated using these images labels using Equation 6.3 is shown in bottom left panel of figures Figures 6.2 and 6.3. In the bottom right panel of the figures we show the gamma corrected output. The gamma used for correction was 2.2. It can be observed that the resultant images present details of all the regions which were depicted individually in the multiple exposed images.

In order to judge the quality of of the dynamic range compression, we use the Dynamic Range Independent Image Quality Assessment metric (DRIMQA) proposed by Aydin et al. [2008]. Due to difference in the dynamic ranges of the original HDR and the resultant LDR images it is difficult to directly compare the input and output images. DRIMQA solves this problem by using a model of the human visual system to define distortions based on the detection and classification of visible changes in the image structure. This metric generates a distortion map that shows the loss of visible features, the amplification of invisible features and reversal of contrast polarity. Thus a quantitative measure for dynamic range compression quality is obtained. Dynamic range compression that involves lower loss of visible contrast, lower amplification of invisible contrast and lower reversal of contrast polarity results in better dynamic range compression quality.

We ran this compression quality metric on our method results and compared them with the results of range of methods widely cited in literature as well as alternatives available commercially (Fattal et al. [2002]; Soft [2008]; Meylan et al. [2007]; Kautz [2008]; Kuang et al. [2007]; Photomatix [1][2018]. In Figure 6.4 we show the images used for the methods comparison. These images have been used widely in literature [2] for dynamic range applications. For each image we took the radiance map and applied our dynamic range compression algorithm and the methods under comparison. We then calculated the DRIMQA metric for each resultant image. In Figures 6.5, 6.6, 6.7and 6.8 we show the comparison based on DRIMQA outcome for each method. The quality assessment is done by calculating the change in contrast in the dynamic range compressed image with respect to the original HDR image. For our comparison, we use dynamic range compressed images without gamma correction which are shown in the resultant images show in the figures. In these figures the error images relate to the DRIMQA metric resultant image depicting change in the contrast. Green regions in the error images refer to the places where loss of contrast took place. Blue regions correspond to places where amplification of contrast took place and red corresponds to the pixels where contrast reversal took place. It can be observed from the figures that our method preserves overall contrast better than other methods in comparison.

In order to judge the results quantitatively, we calculated the percentage of pixels suffering from each change of contrast type. For the comparison of methods, we made table for each change of contrast type. In Table 6.1, we show the result for the loss of contrast for methods under comparison when run on the images shown in Figure 6.4. In Table 6.2, we show the result for the amplification of the contrast for the methods whereas Table 6.3 shows the results for the reversal of the contrast. In the tables, we have written in bold, the outcome of best performing method for each image. At the bottom of each comparison table, we show the mean and standard deviation for all the images under comparison of each method. It can be observed, that no given method stands as the absolute best in all change of contrast types. Our

[1] downloadable at https://www.hdrsoft.com

[2] downloadable at http://www.pauldebevec.com/Research/HDR and http://pfstools.sourceforge.net/hdr_gallery.html

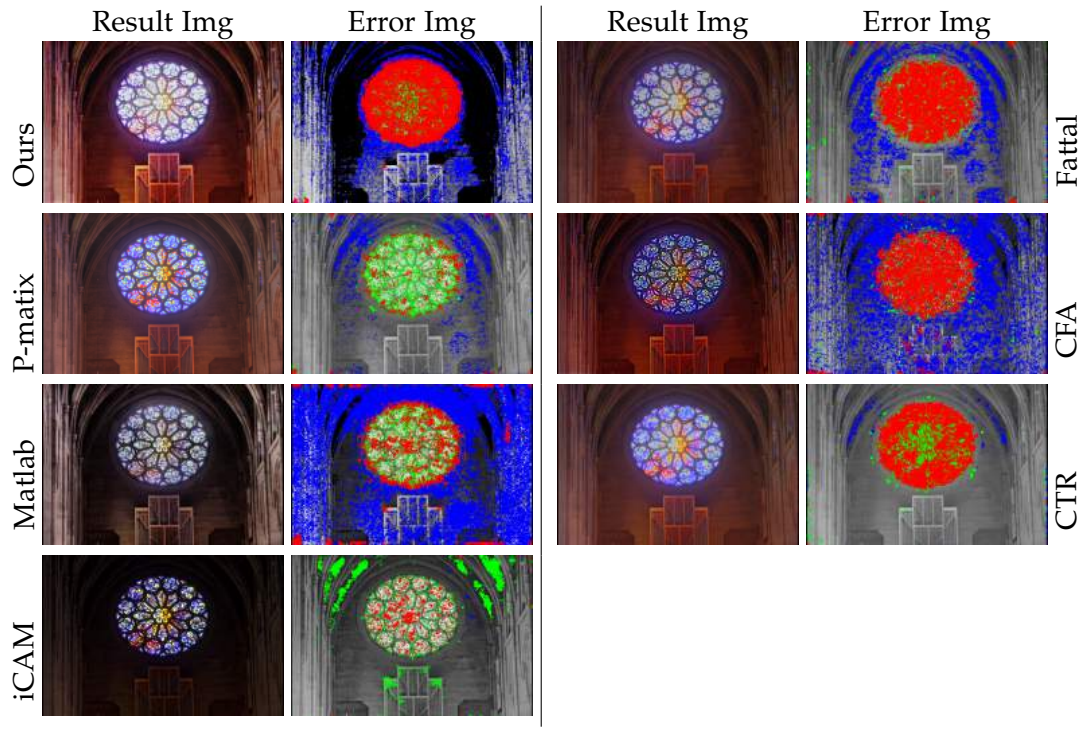

Figure 6.5: Comparison of our method results with the results of methods widely available and cited in literature. For each method the first column contains the gamma corrected output and second column contains the output of the dynamic range independent image quality assessment (DRIMQA) for the respective method

	Methods						
Images	CFA	CTR	Fattal	iCAM	Matlab	Photomatix	Ours
Cathedral	77.164	93.495	97.297	83.066	60.224	79.940	**46.404**
Indoor Room	21.809	0.967	71.337	**0.209**	1.858	0.561	11.389
Memorial Church	10.200	7.569	74.352	16.636	4.920	**2.722**	3.071
Nave	**5.385**	9.996	11.702	6.433	8.268	9.113	11.437
Rosette	**2.090**	4.077	2.393	6.676	4.089	6.049	2.204
mpi atrium 1	32.745	6.102	68.429	2.341	3.347	2.858	**2.118**
mpi atrium 3	22.801	15.829	86.127	19.850	**8.737**	12.665	13.537
nancy church 1	10.376	9.594	90.968	10.442	**5.344**	7.032	10.715
nancy church 2	17.651	11.319	95.316	**4.707**	7.092	8.801	7.296
nancy church 3	29.160	2.050	7.603	**0.791**	3.662	1.448	7.199
Average	22.938	16.100	60.552	15.115	**10.754**	13.119	11.537
Standard Deviation	21.518	27.560	38.121	24.74	17.521	23.791	**12.930**

Table 6.1: Table showing Loss of contrast of method for different images

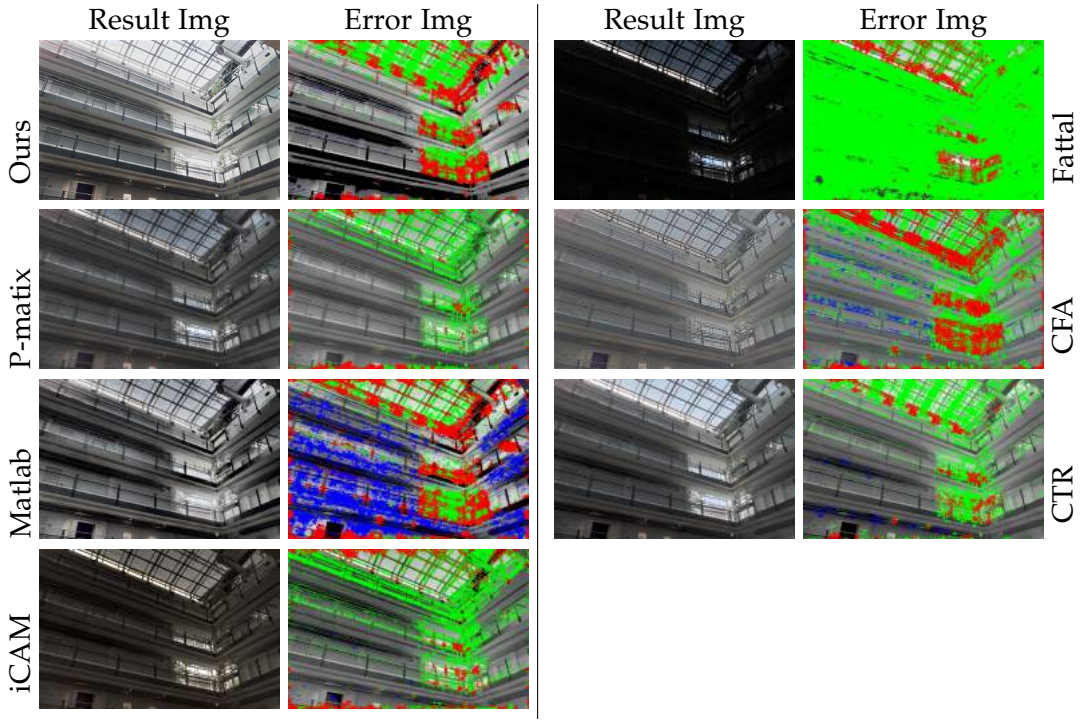

Figure 6.6: Comparison of our method results with the results of methods widely available and cited in literature. For each method the first column contains the gamma corrected output and second column contains the output of the dynamic range independent image quality assessment (DRIMQA) for the respective method

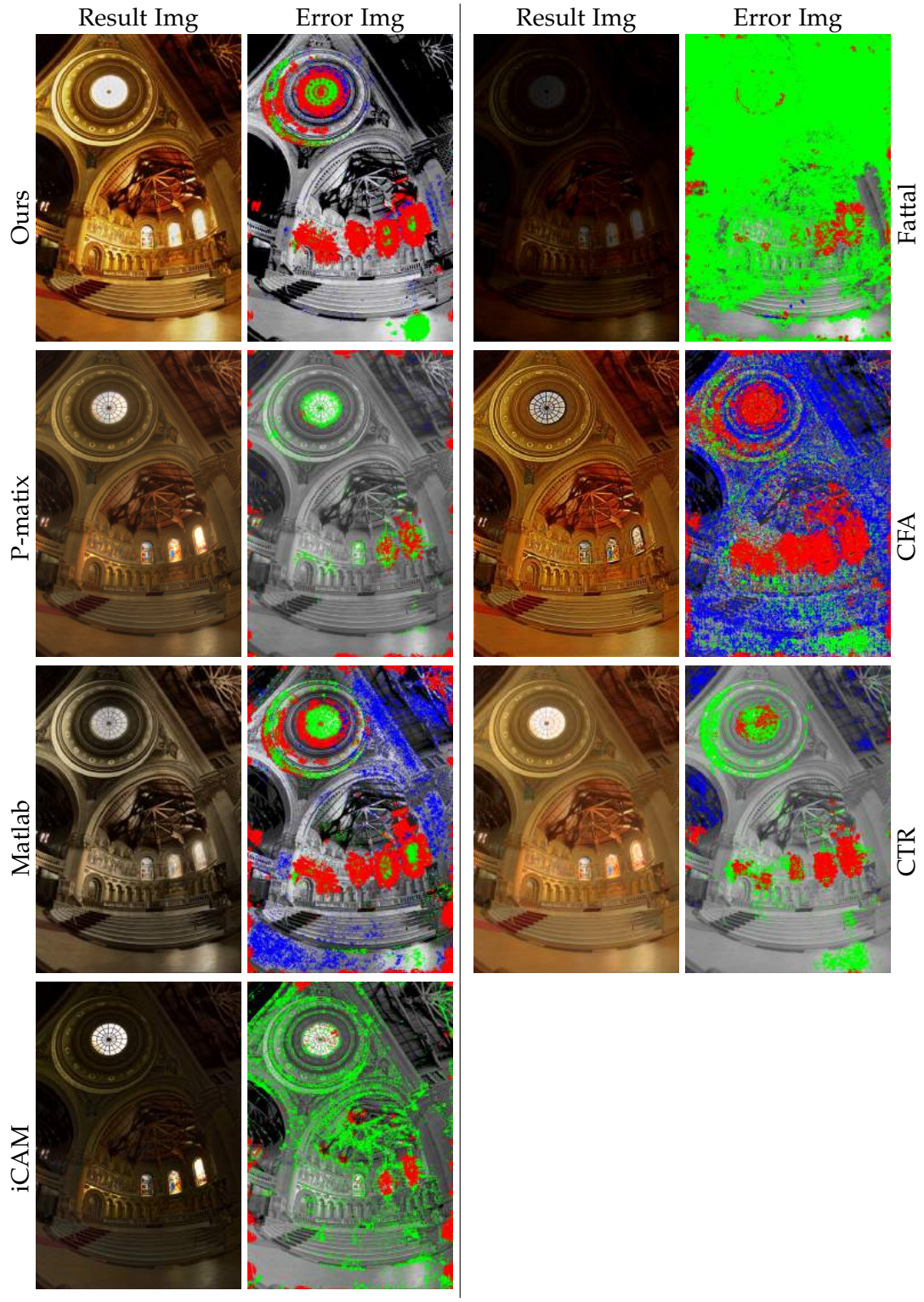

Figure 6.7: Comparison of our method results with the results of methods widely available and cited in literature. For each method the first column contains the gamma corrected output and second column contains the output of the dynamic range independent image quality assessment (DRIMQA) for the respective method

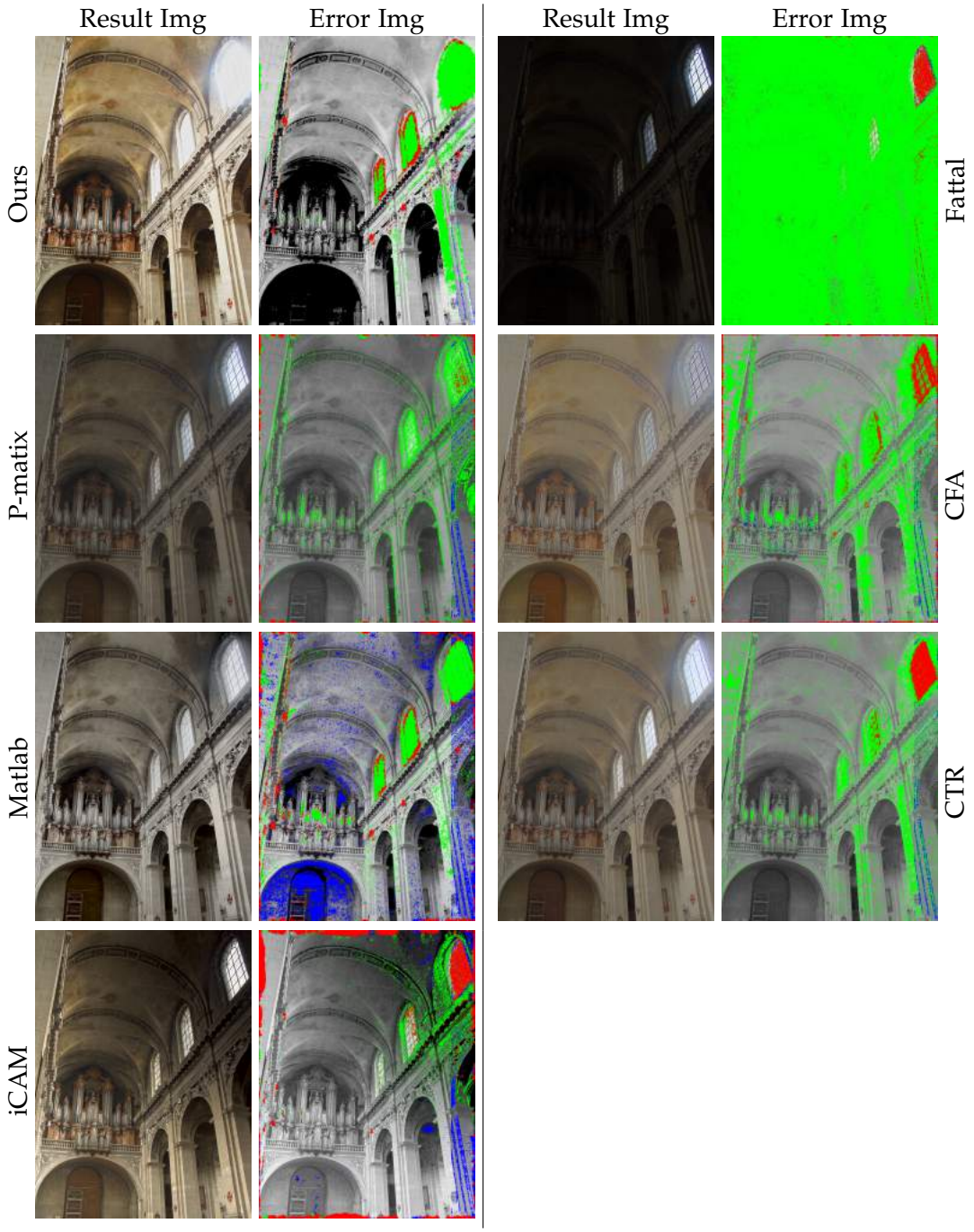

Figure 6.8: Comparison of our method results with the results of methods widely available and cited in literature. For each method the first column contains the gamma corrected output and second column contains the output of the dynamic range independent image quality assessment (DRIMQA) for the respective method

Images	Methods						
	CFA	CTR	Fattal	iCAM	Matlab	Photomatix	Ours
Cathedral	1.010	0.008	0.004	0.161	0.405	**0.003**	0.068
Indoor Room	18.731	0.260	0.322	2.243	7.347	0.427	**0.015**
Memorial Church	32.766	2.146	0.113	**0.079**	11.404	0.291	0.506
Nave	26.743	0.608	0.345	0.234	14.568	0.327	**0.036**
Rosette	33.658	1.236	13.474	**0.286**	46.717	7.054	4.164
mpi atrium 1	1.714	0.883	**0.245**	0.924	11.465	0.320	0.655
mpi atrium 3	3.486	0.695	0.118	0.367	19.877	**0.023**	0.331
nancy church 1	0.422	0.182	**0.111**	0.961	23.553	0.312	0.186
nancy church 2	1.26	0.780	**0.148**	2.504	10.199	1.448	0.386
nancy church 3	3.317	0.242	0.897	2.357	5.807	**0.187**	0.253
Average	12.311	0.704	1.578	1.012	15.134	1.039	**0.660**
Standard Deviation	14.086	**0.630**	4.187	0.983	12.942	2.152	1.249

Table 6.2: Table showing amplification of contrast of method for different images

method tends to have least average amplification of contrast. It also delivers a quite low loss of contrast. It has been observed that the change of contrast by reversal has negligible visual affect in wide range of application (Subramaniam and Biederman [1997]) compared to the change of contrast by amplification or loss. In regards to the consistency of methods performance, our method shows the least standard deviation for loss of contrast and close to the least standard deviation for contrast amplification. Similarly, for the contrast reversal, it exhibits a very low standard deviation.

6.4 Summary

We have presented a method that compresses the dynamic range of HDR images by generating a set of LDR images representing contrast information at different exposures. It uses the contrast information of LDR images based at neighborhood information to minimize a graph regularization energy function. This result in a label set which is used to fuse the LDR images to form the final dynamic range compressed image. We presented the outcome of our method and compared our results with methods widely available in literature and others used professionally. We used a dynamic range independent image quality assessment metric for method comparison and presented a quantitative analysis of our results. Our comparisons

	Methods						
Images	CFA	CTR	Fattal	iCAM	Matlab	Photomatix	Ours
Cathedral	14.180	0.798	2.686	3.842	5.588	**0.485**	1.224
Indoor Room	5.965	1.847	0.568	0.540	0.416	**0.124**	0.305
Memorial Church	18.112	4.518	3.743	3.029	11.552	**2.608**	8.144
Nave	12.036	5.890	5.293	3.098	6.025	**2.045**	6.243
Rosette	13.902	13.285	14.819	**2.669**	7.997	3.014	13.563
mpi atrium 1	30.578	9.469	8.174	5.490	30.399	**3.987**	19.513
mpi atrium 3	14.725	3.604	5.969	5.426	16.267	**2.695**	11.261
nancy church 1	2.930	2.530	1.698	4.935	2.269	**0.558**	1.398
nancy church 2	2.888	1.758	1.507	4.830	3.127	**1.030**	1.453
nancy church 3	67.513	1.623	1.168	3.502	1.980	**0.862**	1.962
Average	18.283	4.532	4.563	3.736	8.562	**1.741**	6.507
Standard Deviation	19.111	4.014	4.346	1.524	9.065	**1.303**	6.523

Table 6.3: Table showing reverse of contrast of method for different images

show an improved and consistent performance of our method.

Chapter 7

Conclusion

7.1 Summary

In this thesis, we have proposed methods to improve the imaging pipeline of digital cameras. Our proposed methods cover three steps of the imaging pipeline (i) automatic exposure control, (ii) radiometric calibration, and (iii) dynamic range compression of radiance map images or high dynamic range images for display on low dynamic range devices.

Our first contribution is aimed at refining the input of the imaging system by automatically adjusting their exposure according to the scene. We proposed two approaches for exposure control. Although our proposed methods focus on spectral camera architectures, our proposed methods are generic in nature and can be extended to any architecture of cameras used. Our first method took the basic equations governing CCD and applied a histogram equalization algorithms to a closed form solution for exposure time calculation. This method tried to estimate a target cumulative distribution function of brightness values per band for its exposure time calculation. We showed results of our method at different wavelength channels. These results included the effects of changing the target cumulative distribution function, and the lighting conditions. It was observed that the method remains largely unaffected by changing initial exposure time for the method.

Our second method, used a regularized non negative least squares optimization technique to estimate the exposure time. Our exposure time calculation was based upoisn a formal stability analysis which assures stability and controllability conditions are satisfied. Our method was also applicable to a variety of architectures of

cameras with different acquisition schemes. We applied our method result on staring array hyperspectral camera, multi-CCD multispectral camera and a single CCD RGB camera, and the results produced were validated by the use of color charts photopic response. For each camera architecture, we showed that our method stretches and linearizes the contrast of spectral power images with our calculated exposure time.

Our second contribution was to propose a method for radiometric calibration of cameras. Our proposed method was devoid of free parameters and can perform radiometric calibration with multiple as well as single images. The method used an Lq Weisfeld algorithm which employs weights to approximate an L1 cost function minimization which effectively reduces the error in estimation of camera response function and radiance map. Our radiometric calibration takes into account channel-wise information and, thus, it can be naturally extended to higher number of channels for spectral cameras. Our results have been quantitatively and qualitatively compared with various methods widely cited in literature. For our comparisons we used the ground truth radiance and camera response functions available in public datasets. We also generated high dynamic range panorama images using our method generated results and presented a comparison of those generated by the alternative methods. The strength of our method is to have least estimation error with minimum number of images and being devoid of parameters settings.

Our third contribution was related to the dynamic range compression of radiance maps or high dynamic range images for low dynamic range display devices. This transformation of incoming scene irradiance into a visible form, is one of the last steps of the imaging pipeline. The method took the input radiance map or high dynamic range image and generated multiple low dynamic range images using exposure selection. It then uses an energy minimization approach to generate probability maps for each LDR image based on their contrast information. We presented our method results as a comparison with a range of methods widely referenced in literature and used professionally. Our comparison was based on the change of contrast information and is relevant to the human visual system perception of contrast. These change of contrast information involves the loss, amplification and reversal of contrast due to dynamic range compression. To measure these change of contrast we

used a dynamic range independent image quality assessment method.

7.2 Future Work

Our work in this thesis has opened up a number of interesting avenues for us to explore in the light of our findings. It would be interesting to look further to other important parameters involved in the imaging pipeline, such as:

- **Estimating the gain of the camera and its effect on the exposure** Gain is the ratio used for converting the number of photoelectrons stored in pixel wells to a digital number. Camera manufacturers offer a similar parameter for users so as to modify the sensitivity of the sensor by changing the camera ISO. In many night photography, surveillance and astronomy applications it is essential to adjust ISO along with the shutter speed. It would be interesting to estimate the sensor gain taking into account the sensor equations and desired ISO setting to better adjust the exposure accordingly.

- **Estimating scene illuminant and its effect on the exposure** To understand scene light conditions is one of the active problem in computer vision and scene analysis. Work has been done extensively to extract illuminant information from images Finlayson and Schaefer [2001a]; Huynh and Robles-Kelly [2010b]. Our CCD equations coupled with physics based illuminant estimation models like the dichromatic reflection model Finlayson and Schaefer [2001b] can be an important area to work in.

- **Estimating color matching functions** Color matching functions are the spectral response to incoming irradiance simulating human eye response to generate red, green and blue channels. In order to have a better estimate of the scene radiance from the image intensities, it warrants to estimate color matching functions, which may further improve the imaging pipeline.

- **Exposure adjustment for different materials** Auto-exposure has a wide set of applications in conjunction with metering techniques. One of the application is to

meter the part of the images containing skin for face detection or human tracking. Most of the modern cameras use the skin tone to better expose humans or faces in the image. Spectral cameras have strong classification capabilities due to their capacity to obtain material spectral signatures. Spectral cameras can be made to adjust the exposure according to materials like human skin in the environment. The work of Huynh and Robles-Kelly [2010a] presents a method for skin recognition with spectral imaging. These methods can efficiently differentiate between a real human and a human picture in the scene. Similarly, in mining applications, the material based exposure can be used to adjust exposure based on mineral composition or terrain features.

- **Video capturing pipeline for spectral cameras** Spectral cameras which have higher spectral resolution often suffer of lower temporal resolution due to the amount of data being processed. Exposure control for rapidly changing scenes for spectral video capture is challenging and needs to be properly visited. The work of Cao et al. [2011] and Du et al. [2009] tries to present different imaging systems to capture multispectral video. These systems can have far ranging applications in tracking, identification and classification as explained in the work of Nguyen et al. [2010]. It is worth exploring how our proposed methods for exposure control can be extended to spectral video capture.

- **Exposure selection** In our method for dynamic range compression we applied a linear exposure selection technique to generate low dynamic range images from the high dynamic range radiance maps. Automatic methods like Pourreza-Shahri and Kehtarnavaz [2015b] Huang and Chiang [2013] Hirakawa and Wolfe [2010] Barakat et al. [2008] have already been proposed in literature. It would be interesting to use statistical models for this estimation.

Appendix A

List of Publications

The Ph.D. candidate has authored the following list of publication relevant to the thesis

- Sohaib, A., Habili, N., & Robles-Kelly, A. (2013, September). Automatic exposure control for multispectral cameras. In Image Processing (ICIP), 2013 20th IEEE International Conference on (pp. 2043-2047). IEEE.
- Sohaib, A., & Robles-Kelly, A. (2015, September). Radiometric calibration for HDR imaging. In Image Processing (ICIP), 2015 IEEE International Conference on (pp. 2815-2819). IEEE.
- Sohaib, A., & Robles-Kelly, A. (2015, November). Exposure Adjustment for Hyperspectral Cameras. In Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on (pp. 1-7). IEEE.
- Sohaib, A., & Robles-Kelly, A. (2015). Exposure time calculation for spectral cameras. Journal of Electronic Imaging, 24(5), 053025-053025.

Bibliography

Aftab, K.; Hartley, R.; and Trumpf, J., 2015. Generalized weiszfeld algorithms for lq optimization. *IEEE transactions on pattern analysis and machine intelligence*, 37, 4 (2015), 728–745. (cited on pages 57, 60, 61, and 79)

Aguerrebere, C.; Almansa, A.; Gousseau, Y.; Delon, J.; and Muse, P., 2014. Single shot high dynamic range imaging using piecewise linear estimators. In *2014 IEEE International Conference on Computational Photography (ICCP)*, 1–10. IEEE. (cited on page 17)

Akbari, H.; Uto, K.; Kosugi, Y.; Kojima, K.; and Tanaka, N., 2011. Cancer detection using infrared hyperspectral imaging. *Cancer science*, 102, 4 (2011), 852–857. (cited on pages 15 and 22)

Ashikhmin, M., 2002. A tone mapping algorithm for high contrast images. In *Proceedings of the 13th Eurographics workshop on Rendering*, 145–156. Eurographics Association. (cited on page 19)

Aydin, T. O.; Mantiuk, R.; Myszkowski, K.; and Seidel, H.-P., 2008. Dynamic range independent image quality assessment. In *ACM Transactions on Graphics (TOG)*, vol. 27, 69. ACM. (cited on page 92)

Banterle, F.; Artusi, A.; Debattista, K.; and Chalmers, A., 2011. *Advanced High Dynamic Range Imaging: Theory and Practice*. CRC Press. (cited on pages 4 and 16)

Banterle, F.; Ledda, P.; Debattista, K.; Chalmers, A.; and Bloj, M., 2007. A framework for inverse tone mapping. *The Visual Computer*, 23, 7 (2007), 467–478. (cited on page 16)

Barakat, N.; Hone, A. N.; and Darcie, T. E., 2008. Minimal-bracketing sets for

high-dynamic-range image capture. *Image Processing, IEEE Transactions on*, 17, 10 (2008), 1864–1875. (cited on pages 88 and 104)

Bassiou, N. and Kotropoulos, C., 2007. Color image histogram equalization by absolute discounting back-off. *Computer Vision and Image Understanding*, 107, 1 (2007), 108–122. (cited on page 46)

Bell, C. S.; Tomaszewski, E. P.; Hansen, A. E.; and Raj, K., 2002. Determining a final exposure serring automatically for a solid state camera without a separate light metering circuit. U.S. Patent US6486915. (cited on page 14)

Boykov, Y. and Jolly, M.-P., 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. In *Intl. Conf. on Computer Vision*, 105–112. (cited on page 84)

Brelstaff, G. J.; Parraga, A.; Troscianko, T.; and Carr, D., 1995. Hyperspectral camera system: acquisition and analysis. *Geographic Information Systems, Photogrammetry, and Geological/Geophysical Remote Sensing*, 150 (1995). (cited on page 15)

Brucalassi, A.; Araiza-Durán, J.; Pignata, G.; Campana, S.; Claudi, R.; Schipani, P.; Aliverti, M.; Baruffolo, A.; Ben-Ami, S.; Biondi, F.; et al., 2018. The acquisition camera system for soxs at ntt. In *Ground-based and Airborne Instrumentation for Astronomy VII*, vol. 10702, 107022M. International Society for Optics and Photonics. (cited on page 15)

Burton, K. B.; Owsley, C.; and Sloane, M. E., 1992. Aging and neural spatial contrast sensitivity: Photopic vision. (1992). (cited on page 38)

Cao, X.; Tong, X.; Dai, Q.; and Lin, S., 2011. High resolution multispectral video capture with a hybrid camera system. *IEEE'Conference'on'Computer'Vision'and'Pattern'Recognition*, (2011). (cited on page 104)

Caras, T. and Karnieli, A., 2015. Ground-level classification of a coral reef using a

hyperspectral camera. *Remote Sensing*, 7, 6 (2015), 7521–7544. (cited on pages 15 and 22)

Chi, C.; Yoo, H.; and Ben-Ezra, M., 2010. Multi-spectral imaging by optimized wide band illumination. *Int. Journal of Computer Vision*, 86 (2010), 140ï£¡151. (cited on pages 15 and 22)

Chiu, K.; Herf, M.; Shirley, P.; Swamy, S.; Wang, C.; Zimmerman, K.; et al., 1993. Spatially nonuniform scaling functions for high contrast images. In *Graphics Interface*, 245–245. CANADIAN INFORMATION PROCESSING SOCIETY. (cited on page 19)

Chung, F., 1997. *Spectral Graph Theory*. American Mathematical Society. (cited on page 87)

CIE, 1932. *Commission Internationale de lï£¡Eï£¡clairage Proceedings, 1931*. Cambridge University Press. (cited on pages xiii, 22, and 28)

Cleveland, W. S., 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74, 368 (1979), 829–836. (cited on page 64)

Dale, L. M.; Thewis, A.; Boudry, C.; Rotar, I.; Dardenne, P.; Baeten, V.; and Pierna, J. A. F., 2013. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Applied Spectroscopy Reviews*, 48, 2 (2013), 142–159. doi:10.1080/05704928.2012.705800. http://dx.doi.org/10.1080/05704928.2012.705800. (cited on pages 15 and 22)

Davis, T., 2006. *Direct Methods for Sparse Linear Systems*. SIAM. (cited on page 87)

Debevec, P. E. and Malik, J., 1997. Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, 369–378. (cited on pages xiii, 1, 2, 10, 17, and 65)

Debevec, P. E. and Malik, J., 2008. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, 31. ACM. (cited on page 18)

DeBoer, D. R.; Gough, R. G.; Bunton, J. D.; Cornwell, T. J.; Beresford, R. J.; Johnston, S.; Feain, I. J.; Schinckel, A. E.; Jackson, C. A.; Kesteven, M. J.; et al., 2009. Australian ska pathfinder: A high-dynamic range wide-field of view survey telescope. *Proceedings of the IEEE*, 97, 8 (2009), 1507–1521. (cited on page 18)

Dogan, Y.; Cinar, A.; and Ozbay, E., 2017. Correction of real-time exposure in hyperspectral bands for industrial cameras. (2017). (cited on page 15)

Drago, F.; Myszkowski, K.; Annen, T.; and Chiba, N., 2003. Adaptive logarithmic mapping for displaying high contrast scenes. In *Computer Graphics Forum*, vol. 22, 419–426. Wiley Online Library. (cited on page 19)

Du, H.; Tong, X.; Cao, X.; and Lin, S., 2009. A prism-based system for multispectral video acquisition. *In Proc.ICCV*, (2009). (cited on page 104)

Durand, F. and Dorsey, J., 2002. Fast bilateral filtering for the display of high-dynamic-range images. In *ACM transactions on graphics (TOG)*, vol. 21, 257–266. ACM. (cited on page 19)

Dwight, J. G.; Tkaczyk, T. S.; Alexander, D.; Pawlowski, M. E.; Luvall, J. C.; Tatum, P. F.; and Jedlovec, G. J., 2018. Compact snapshot image mapping spectrometer (snap-ims) for hyperspectral data cube acquisition using unmanned aerial vehicle (uav) environmental imaging. In *Next-Generation Spectroscopic Technologies XI*, vol. 10657, 106570G. International Society for Optics and Photonics. (cited on page 15)

Eilertsen, G.; Mantiuk, R. K.; and Unger, J., 2017. A comparative review of tone-mapping algorithms for high dynamic range video. In *Computer Graphics Forum*, vol. 36, 565–592. Wiley Online Library. (cited on page 16)

Fattal, R.; Lischinski, D.; and Werman, M., 2002. Gradient domain high dynamic

range compression. In *ACM Transactions on Graphics (TOG)*, vol. 21, 249–256. ACM. (cited on pages 19 and 93)

Ferwerda, J. A.; Pattanaik, S. N.; Shirley, P.; and Greenberg, D. P., 1996. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 249–258. ACM. (cited on page 19)

Finlayson, G. D. and Schaefer, G., 2001a. Convex and non-convex illuminant constraints for dichromatic colour constancy. In *IEEE Conf. on Computer Vision and Pattern Recognition*, I:598–604. (cited on page 103)

Finlayson, G. D. and Schaefer, G., 2001b. Solving for colour constancy using a constrained dichromatic reflection model. *International Journal of Computer Vision*, 42, 3 (2001), 127–144. (cited on page 103)

Foster, D. H.; Amano, K.; Nascimento, S. M. C.; and Foster, M. J., 2006. Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A*, 23, 10 (2006), 2359–2372. (cited on pages xiii, 22, and 65)

Garrido-Novell, C.; Pérez-Marin, D.; Amigo, J. M.; Fernández-Novales, J.; Guerrero, J. E.; and Garrido-Varo, A., 2012. Grading and color evolution of apples using rgb and hyperspectral imaging vision cameras. *Journal of Food Engineering*, 113, 2 (2012), 281–288. (cited on pages 15 and 22)

Grady, L., 2006. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2006). (cited on page 86)

Groezinger, J. L., 2000. Method and system for remotely viewing and configuring output from a medical imaging device. US Patent 6,101,407. (cited on page 1)

Hasinoff, S. W.; Durand, F.; and Freeman, W. T., 2010. Noise-optimal capture for high dynamic range photography. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 553–560. IEEE. (cited on page 18)

Hirakawa, K. and Wolfe, P. J., 2010. Optimal exposure control for high dynamic range imaging. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, 3137–3140. IEEE. (cited on pages 88 and 104)

Hirokazu, A. and Tadazumi, S., 1975. Control circuit for an exposure meter. US Patent 3,879,137. (cited on page 14)

Huang, K.-F. and Chiang, J.-C., 2013. Intelligent exposure determination for high quality hdr image generation. In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, 3201–3205. IEEE. (cited on pages 88 and 104)

Huynh, C. P. and Robles-Kelly, A., 2010a. Hyperspectral imaging for skin recognition and biometrics. In *International Conference on Image Processing*. (cited on page 104)

Huynh, C. P. and Robles-Kelly, A., 2010b. A solution of the dichromatic model for multispectral photometric invariance. *International Journal of Computer Vision*, 90, 1 (2010), 1–27. (cited on page 103)

Ikeda, E., 1998. Image data processing apparatus for processing combined image signals in order to extend dynamic range. US Patent 5,801,773. (cited on page 16)

Janesick, J. R., 2001. *Scientific charge-coupled devices*, vol. 117. SPIE press Bellingham, Washington. (cited on page 12)

Jiang, J.; Liu, D.; Gu, J.; and Süsstrunk, S., 2013. What is the space of spectral sensitivity functions for digital color cameras? In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, 168–179. IEEE. (cited on page 62)

Jinno, T. and Okuda, M., 2012. Multiple exposure fusion for high dynamic range image acquisition. *IEEE Transactions on Image Processing*, 21, 1 (2012), 358–365. (cited on pages 17 and 18)

Johnson, B. K., 1984. Photographic exposure control system and method. U.S. Patent US4423936. (cited on page 14)

Johnson, G. M. and Fairchild, M. D., 2003. Rendering hdr images. In *Color and Imaging Conference*, vol. 2003, 36–41. Society for Imaging Science and Technology. (cited on page 18)

Jung, J. I. and Ho, Y. S., 2013. Mrf-based high dynamic range image generation. In *Digital Signal Processing*, 1–4. (cited on page 16)

Kalantari, N. K. and Ramamoorthi, R., 2017. Deep high dynamic range imaging of dynamic scenes. *ACM Trans. Graph*, 36, 4 (2017), 144. (cited on page 16)

Katsuhiko, O., 2010. Modern control engineering. (2010). (cited on pages 41 and 44)

Kaur, M.; Kaur, J.; and Kaur, J., 2011. Survey of contrast enhancement techniques based on histogram equalization. *International Journal of Advanced Computer Science and Applications*, 2, 7 (2011). (cited on page 26)

Kautz, M. H. K. J., 2008. Consistent tone reproduction. (2008). (cited on pages 19 and 93)

Khan, Z.; Mian, A.; and Hu, Y., 2011. Contour code: Robust and efficient multispectral palmprint encoding for human recognition. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, 1935–1942. IEEE. (cited on pages 15 and 22)

Klinger, T., 2003. *Image processing with LabVIEW and IMAQ Vision*. Prentice Hall Professional. (cited on page 12)

Kremens, R.; Sampat, N.; Venkataraman, S.; and Yeh, T., 1999. System implications of implementing auto-exposure on consumer digital cameras. In *Proceedings of the SPIE Electronic Imaging*. (cited on page 14)

Kuang, J.; Johnson, G. M.; and Fairchild, M. D., 2007. icam06: A refined image appearance model for hdr image rendering. *Journal of Visual Communication and Image Representation*, 18, 5 (2007), 406–414. (cited on page 93)

Kubota, M.; Kato, T.; Suzuki, S.; Maruyama, H.; Shidara, K.; Tanioka, K.; Sameshima, K.; Makishima, T.; Tsuji, K.; Hirai, T.; et al., 1996. Ultrahigh-

sensitivity new super-harp camera. *IEEE transactions on broadcasting*, 42, 3 (1996), 251–258. (cited on page 1)

Kulasekara, S. I.; Patel, S. R.; Shahidi, A. M.; Flanagan, J. G.; and Hudson, C., 2014. Detection of changes in retinal vessel blood oxygen saturation using hyperspectral imaging. *Investigative Ophthalmology & Visual Science*, 55, 13 (2014), 199–199. (cited on pages 15 and 22)

Kuno, T.; Sugiura, H.; and Matoba, N., 1998. A new automatic exposure system for digital still cameras. *Consumer Electronics, IEEE Transactions on*, 44, 1 (1998), 192–199. (cited on page 14)

Lang, K.; Madhavan, V.; Hoffman, J.; Hudson, E.; Eisaki, H.; Uchida, S.; and Davis, J., 2002. Imaging the granular structure of high-tc superconductivity in underdoped bi2sr2cacu2o8+ δ. *Nature*, 415, 6870 (2002), 412–416. (cited on page 1)

Larson, G. W.; Rushmeier, H.; and Piatko, C., 1997. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3, 4 (1997), 291–306. (cited on pages 19 and 78)

Lee, J.-S.; Jung, Y.-Y.; Kim, B.-S.; and Ko, S.-J., 2001. An advanced video camera system with robust af, ae, and awb control. *Consumer Electronics, IEEE Transactions on*, 47, 3 (2001), 694–699. (cited on page 14)

Lieb, D. F. and Russell, A. I., 2014. High dynamic range display systems. US Patent 8,705,133. (cited on page 18)

Lin, S.; Gu, J.; Yamazaki, S.; and Shum, H.-Y., 2004. Radiometric calibration from a single image. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, II–938. IEEE. (cited on page 17)

Lin, S. and Zhang, L., 2005. Determining the radiometric response function from a single grayscale image. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, 66–73. IEEE. (cited on page 17)

Liu, Z.; Yang, D.; and Zhu, M., 2011. Comparison of hdr image compression algorithms. *Imaging Science and Photochemistry*, 29, 1 (2011), 32–44. (cited on page 19)

Long, J. and Li, S., 2011. Near infrared face image quality assessment system of video sequences. In *Image and Graphics (ICIG), 2011 Sixth International Conference on*, 275–279. IEEE. (cited on pages 15 and 22)

Malbet, F.; Yu, J.; and Shao, M., 1995. High-dynamic-range imaging using a deformable mirror for space coronography. *Publications of the Astronomical Society of the Pacific*, (1995), 386–398. (cited on page 18)

Martin, D. C.; Fanson, J.; Schiminovich, D.; Morrissey, P.; Friedman, P. G.; Barlow, T. A.; Conrow, T.; Grange, R.; Jelinsky, P. N.; Milliard, B.; et al., 2005. The galaxy evolution explorer: A space ultraviolet survey mission. *The Astrophysical Journal Letters*, 619, 1 (2005), L1. (cited on page 1)

Matsushita, Y. and Lin, S., 2007. Radiometric calibration from noise distributions. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8. IEEE. (cited on page 17)

McCamy, C. S.; Marcus, H.; and Davidson, J. G., 1976. A color rendition charts. *Journal of Applied Photographic Engineering*, 11(3) (1976), 95–99. (cited on page 47)

McCrindle, I. J.; Grant, J.; Drysdale, T. D.; and Cumming, D. R., 2014. Multispectral materials: hybridisation of optical plasmonic filters and a terahertz metamaterial absorber. *Advanced Optical Materials*, 2, 2 (2014), 149–153. (cited on page 1)

Meylan, L.; Alleysson, D.; and Süsstrunk, S., 2007. Model of retinal local adaptation for the tone mapping of color filter array images. *JOSA A*, 24, 9 (2007), 2807–2816. (cited on page 93)

Mitsunaga, T. and Nayar, S. K., 1999. Radiometric self calibration. In *Computer Vision and Pattern Recognition*, vol. 1. (cited on pages 4, 16, 17, 18, and 65)

Nagatani, K.; Akiyama, K.; Yamauchi, G.; Otsuka, H.; Nakamura, T.; Kiribayashi, S.; Yoshida, K.; Hada, Y.; Fujino, K.; Izu, T.; et al., 2013. Volcanic ash observation in active volcano areas using teleoperated mobile robots-introduction to our robotic-volcano-observation project and field experiments. In *2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 1–6. IEEE. (cited on page 1)

Nagendra, H.; Lucas, R.; Honrado, J. P.; Jongman, R. H.; Tarantino, C.; Adamo, M.; and Mairota, P., 2013. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators*, 33 (2013), 45–59. (cited on page 1)

Nayar, S. K. and Mitsunaga, T., 2000. High dynamic range imaging: Spatially varying pixel exposures. In *Computer Vision and Pattern Recognition*, vol. 1, 472–479. (cited on page 16)

Nelson, N. R., 1966. Light metering means for cameras. US Patent 3,282,178. (cited on page 14)

Nevalainen, N. and Pellinen, T., 2016. The use of a thermal camera for quality assurance of asphalt pavement construction. *International Journal of Pavement Engineering*, 17, 7 (2016), 626–636. (cited on page 1)

Nguyen, H. V.; Banerjee, A.; and Chellappa, R., 2010. Tracking via object reflectance using a hyperspectral video camera. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, 44 –51. (cited on page 104)

Nguyen, R.; Prasad, D. K.; and Brown, M. S., 2014. Raw-to-raw: Mapping between image sensor color responses. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3398–3405. (cited on page 13)

Norwood, D. W., 1940. Exposure meter. US Patent 2,214,283. (cited on page 14)

Nourani-Vatani, N. and Roberts, J., 2007. Automatic camera exposure control. In

Proceeding of the Australasian Conference on Robotics and Automation. (cited on page 14)

Pajares, G., 2015. Overview and current status of remote sensing applications based on unmanned aerial vehicles (uavs). *Photogrammetric Engineering & Remote Sensing*, 81, 4 (2015), 281–329. (cited on page 1)

Pattanaik, S. N.; Ferwerda, J. A.; Fairchild, M. D.; and Greenberg, D. P., 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 287–298. ACM. (cited on page 19)

Pattanaik, S. N.; Tumblin, J.; Yee, H.; and Greenberg, D. P., 2000. Time-dependent visual adaptation for fast realistic image display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 47–54. ACM Press/Addison-Wesley Publishing Co. (cited on page 19)

Pichette, J.; Charle, W.; and Lambrechts, A., 2017. Fast and compact internal scanning cmos-based hyperspectral camera: the snapscan. In *Photonic Instrumentation Engineering IV*, vol. 10110, 1011014. International Society for Optics and Photonics. (cited on page 15)

Pourreza-Shahri, R. and Kehtarnavaz, N., 2015a. Automatic exposure selection for high dynamic range photography. In *Consumer Electronics (ICCE), 2015 IEEE International Conference on*, 471–472. IEEE. (cited on page 88)

Pourreza-Shahri, R. and Kehtarnavaz, N., 2015b. Automatic exposure selection for high dynamic range photography. In *Consumer Electronics (ICCE), 2015 IEEE International Conference on*, 471–472. IEEE. (cited on page 104)

Reinhard, E. and Devlin, K., 2005. Dynamic range reduction inspired by photoreceptor physiology. *Visualization and Computer Graphics, IEEE Transactions on*, 11, 1 (2005), 13–24. (cited on page 19)

Reinhard, E.; Heidrich, W.; Debevec, P.; Pattanaik, S.; Ward, G.; and

Myszkowski, K., 2010. *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann. (cited on page 18)

Riley, D. N. and Hecker, C. A., 2013. Mineral mapping with airborne hyperspectral thermal infrared remote sensing at cuprite, nevada, usa. In *Thermal Infrared Remote Sensing*, 495–514. Springer. (cited on page 1)

Robertson, M. A.; Borman, S.; and Stevenson, R. L., 2003. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging*, 12, 2 (2003), 219–228. (cited on pages 17, 18, 65, and 79)

Robles-Kelly, A. and Huynh, C. P., 2012a. *Imaging spectroscopy for scene analysis*. Springer Science & Business Media. (cited on page 10)

Robles-Kelly, A. and Huynh, C. P., 2012b. *Imaging spectroscopy for scene analysis*. Springer Science & Business Media. (cited on page 22)

Robles-Kelly, A. and Huynh, C. P., 2013. *Imaging Spectroscopy for Scene Analysis*. Springer. (cited on page 59)

Rogers, T. E. and Cope, J. M., 2010. System and method for camera metering based on flesh tone detection. U.S. Patent US7847830 B2. (cited on page 15)

Rohs, M. and Gfeller, B., 2004. *Using camera-equipped mobile phones for interacting with real-world objects*. na. (cited on page 1)

Schechner, Y.; Nayar, S.; and Belhumeur, P. N., 2003. A theory of multiplexed illumination. In *Int. Conference on Computer Vision*, II:808–816. (cited on pages 15 and 22)

Seetzen, H.; Heidrich, W.; Stuerzlinger, W.; Ward, G.; Whitehead, L.; Trentacoste, M.; Ghosh, A.; and Vorozcovs, A., 2004. High dynamic range display systems. In *ACM Transactions on Graphics (TOG)*, vol. 23, 760–768. ACM. (cited on page 18)

Shimizu, S.; Kondo, T.; Kohashi, T.; Tsuruta, M.; and Komuro, T., 1992. A new algorithm for exposure control based on fuzzy logic for video cameras. *IEEE Transactions on Consumer Electronics*, 38 (1992), 617–ï£¡623. (cited on page 14)

Soft, H., 2008. Photomatix. (cited on page 93)

Sohaib, A. and Robles-Kelly, A., 2015. Radiometric calibration for hdr imaging. In *Image Processing (ICIP), 2015 IEEE International Conference on*, 2815–2819. IEEE. (cited on page 18)

Spivak, A.; Belenky, A.; Fish, A.; and Yadid-Pecht, O., 2009. Wide-dynamic-range cmos image sensorsâĂŤcomparative performance analysis. *Electron Devices, IEEE Transactions on*, 56, 11 (2009), 2446–2461. (cited on page 18)

Stark, P. B. and Parker, R. L., 1995. Bounded-variable least-squares: an algorithm and applications. *Computational Statistics*, 10 (1995), 129–129. (cited on page 46)

Stiles, W. S. and Burch, J. M., 1955. Interim report to the Commission Internationale de l'Éclairage Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching. *Optica Acta*, 2 (1955), 168–181. (cited on pages 36 and 37)

Subramaniam, S. and Biederman, I., 1997. Does contrast reversal affect object identification? In *Investigative Ophthalmology & Visual Science*, vol. 38, 4638–4638. LIPPINCOTT-RAVEN PUBL 227 EAST WASHINGTON SQ, PHILADELPHIA, PA 19106. (cited on page 98)

Sudhama, A.; Cutone, M. D.; Hou, Y.; Goel, J.; Stolitzka, D.; Jacobson, N.; Allison, R. S.; and Wilcox, L. M., 2018. 85-1: Visually lossless compression of high dynamic range images: A large-scale evaluation. *SID Symposium Digest of Technical Papers*, 49, 1 (2018), 1151–1154. (cited on page 19)

Sun, N.; Mansour, H.; and Ward, R., 2010. Hdr image construction from multi-exposed stereo ldr images. In *IEEE International Conference on Image Processing*, 2973–2976. (cited on page 17)

Sung, M. C.; Wang, T. H.; and Lien, J. J., 2007. High dynamic range scene realization using two complementary images. In *Asian Conference on Computer Vision*, 261–270. (cited on page 16)

Takagi, T., 1997. Auto-exposure device of a camera. U.S. Patent US5596387. (cited on page 14)

Thenkabail, P. S.; Lyon, J. G.; and Huete, A., 2016. *Hyperspectral remote sensing of vegetation*. CRC Press. (cited on page 1)

Tomasi, C. and Manduchi, R., 1998. Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on*, 839–846. IEEE. (cited on page 19)

Tumblin, J. and Rushmeier, H., 1993. Tone reproduction for realistic images. *Computer Graphics and Applications, IEEE*, 13, 6 (1993), 42–48. (cited on pages 18 and 19)

Tumblin, J. and Turk, G., 1999. Lcis: A boundary hierarchy for detail-preserving contrast reduction. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 83–90. ACM Press/Addison-Wesley Publishing Co. (cited on page 19)

Tumblin, J. E. J. and Rushmeier, H. E., 1991. Tone reproduction for realistic computer generated images. (1991). (cited on page 18)

Wandell, B. A., 1987. The synthesis and analysis of color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9, 1 (1987), 2–13. (cited on pages 15 and 22)

Ward, G. J., 1994. The radiance lighting simulation and rendering system. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, 459–472. ACM. (cited on page 19)

Weisstein, E. W., 2002. Least squares fitting–polynomial. (2002). (cited on page 64)

Wu, D. and Sun, D.-W., 2013. Colour measurements by computer vision for food quality control–a review. *Trends in Food Science & Technology*, 29, 1 (2013), 5–20. (cited on page 1)

Yadid-Pecht, O., 1999. Wide-dynamic-range sensors. *Optical Engineering*, 38, 10 (1999), 1650–1660. (cited on page 18)

Yang, I.-C.; Hsieh, K.-W.; Tsai, C.-Y.; Huang, Y.-I.; Chen, Y.-L.; and Chen, S., 2014. Development of an automation system for greenhouse seedling production management using radio-frequency-identification and local remote sensing techniques. *Engineering in Agriculture, Environment and Food*, 7, 1 (2014), 52–58. (cited on page 1)

Zhang, Z., 2012. Microsoft kinect sensor and its effect. *IEEE multimedia*, 19, 2 (2012), 4–10. (cited on page 1)

Zhao, Y. and Liu, N., 2010. Tensor bidirectional reflectance distribution function model graph based multi-spectral face recognition. In *Image and Signal Processing (CISP), 2010 3rd International Congress on*, vol. 4, 1956–1960. IEEE. (cited on pages 15 and 22)